

# The Markoff Group of Transformations in Prime and Composite Moduli

Chen Meiri and Doron Puder  
with an Appendix by Dan Carmon

February 28, 2017

## Abstract

The Markoff group of transformations is a group  $\Gamma$  of affine integral morphisms, which is known to act transitively on the set of all positive integer solutions to the equation  $x^2 + y^2 + z^2 = xyz$ . The fundamental strong approximation conjecture for the Markoff equation states that for every prime  $p$ , the group  $\Gamma$  acts transitively on the set  $X^*(p)$  of non-zero solutions to the same equation over  $\mathbb{Z}/p\mathbb{Z}$ . Recently, Bourgain, Gamburd and Sarnak proved this conjecture for all primes outside a small exceptional set.

In the current paper, we study a group of permutations obtained by the action of  $\Gamma$  on  $X^*(p)$ , and show that for most primes, it is the full symmetric or alternating group. We use this result to deduce that  $\Gamma$  acts transitively also on the set of non-zero solutions in a big class of composite moduli.

Our result is also related to a well-known theorem of Gilman, stating that for any finite non-abelian simple group  $G$  and  $r \geq 3$ , the group  $\text{Aut}(F_r)$  acts on at least one “ $T_r$ -system” of  $G$  as the alternating or symmetric group. In this language, our main result translates to that for most primes  $p$ , the group  $\text{Aut}(F_2)$  acts on a particular  $T_2$ -system of  $\text{PSL}(2, p)$  as the alternating or symmetric group.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Preliminaries</b>	<b>5</b>
<b>3</b>	<b>Alternating Group for <math>p \equiv 1 \pmod{4}</math></b>	<b>7</b>
<b>4</b>	<b>Alternating Group for <math>p \equiv 3 \pmod{4}</math></b>	<b>8</b>
4.1	Deducing Alternating group from primitivity . . . . .	9
4.2	Primitivity for density-1 of the primes . . . . .	10
4.3	Ruling out correlation between two long $\text{rot}_1$ -cycles with same first coordinate .	12
<b>5</b>	<b>Strong Approximation for Square Free Composite Moduli</b>	<b>17</b>
5.1	Transitivity modulo square free composite moduli, assuming alternating groups	17
5.2	Transitivity modulo square free composite moduli, assuming primitivity . . . .	20
<b>6</b>	<b><math>T_2</math>-systems</b>	<b>21</b>
<b>A</b>	<b>On the order of a quadratic integer modulo most primes</b>	
	By Dan Carmon	<b>23</b>

# 1 Introduction

The Markoff surface  $\mathbb{X}$  is the affine surface in  $\mathbb{A}^3$  defined by the equation<sup>1</sup>

$$x^2 + y^2 + z^2 = xyz. \quad (1)$$

The Markoff triples  $\mathcal{M}$  is the set of positive integer solutions to Equation (1), such as  $(3, 3, 3)$ .  $\mathcal{M}$   
The Markoff group of automorphisms of  $\mathbb{X}$  is the group  $\Gamma$  generated by permutations of the  $\Gamma$   
coordinates and the Vieta involutions  $R_1, R_2$  and  $R_3$  where  $R_3(x, y, z) = (x, y, xy - z)$  and  $R_i$   
 $R_1$  and  $R_2$  are defined analogously. It is easy to see that  $\mathcal{M}$  is invariant under  $\Gamma$  and Markoff  
proved that  $\Gamma$  acts transitively on  $\mathcal{M}$  [Mar79, Mar80]. Let  $\Delta$  be the group generated by  
 $\Gamma$  and the involutions that replace two of the coordinates by their negatives. Then the set  
 $\mathbb{X}(\mathbb{Z})$  of integer solutions to (1) has two  $\Delta$ -orbits:  $\{(0, 0, 0)\}$  and its complement  $X^*(\mathbb{Z}) \stackrel{\text{def}}{=} \mathbb{X}(\mathbb{Z}) \setminus \{(0, 0, 0)\}$ .

## Prime Moduli

If  $p$  is a prime number, then  $\mathbb{X}(\mathbb{Z}/p\mathbb{Z})$  is the finite set of solutions to (1) in  $\mathbb{Z}/p\mathbb{Z}$ , and we denote  $X^*(p) = \mathbb{X}(\mathbb{Z}/p\mathbb{Z}) \setminus \{(0, 0, 0)\}$ . The strong approximation conjecture for the Markoff equation  $X^*(p)$   
(1) states that for every prime  $p$ , the reduction mod  $p$  of the set of Markoff triples  $\mathcal{M} \rightarrow X^*(p)$   
is onto. This is clearly equivalent to  $\Gamma$  acting transitively on  $X^*(p)$ . Recently, Bourgain,  
Gamburd and Sarnak proved this conjecture for all primes outside of a small exceptional set:

**Theorem 1.1** (Bourgain-Gamburd-Sarnak [BGS17]). *Let  $E$  be the set of primes for which  $\Gamma$  does not act transitively on  $X^*(p)$ . For any  $\varepsilon > 0$ , the number of primes  $p \leq T$  with  $p \in E$  is at most  $T^\varepsilon$ , for  $T$  large enough.*

*Moreover, for any  $\varepsilon > 0$ , the largest  $\Gamma$ -orbit in  $X^*(p)$  is of size at least  $|X^*(p)| - p^\varepsilon$ , for  $p$  large enough (note that  $|X^*(p)| \sim p^2$ ).*

Let  $\Gamma_p$  be the finite permutation group induced by the action of  $\Gamma$  on  $X^*(p)$ . In the current  $\Gamma_p$   
work we study the nature of this group. The first step here is to notice that  $\Gamma_p$  preserves a  
block structure as follows:

For  $(x, y, z) \in X^*(p)$  denote by  $[x, y, z]$  the block of all solutions obtained from  $(x, y, z)$  by  $[x, y, z]$   
sign changes, so

$$[x, y, z] \stackrel{\text{def}}{=} \{(x, y, z), (x, -y, -z), (-x, y, -z), (-x, -y, z)\}.$$

Then  $\Gamma_p$  preserves this block structure. Let  $Y^*(p)$  denote the set of blocks in  $X^*(p)$ , and  $Y^*(p)$   
 $Q_p$  denote the permutation group induced by the action of  $\Gamma$  (or  $\Gamma_p$ ) on  $Y^*(p)$ . Simulations  $Q_p$   
suggest the following conjecture:

**Conjecture 1.2.** *For every  $p \geq 5$ , the permutation group  $Q_p$  is the full alternating or symmetric group.*

This conjecture was also raised, independently, in [CGMP16, Conjecture 1.3], where the authors also state precisely for which primes one can expect the alternating group ( $p \equiv 3 \pmod{16}$ ) and for which the full symmetric group ( $p \not\equiv 3 \pmod{16}$ ). If this conjecture holds, then roughly speaking,  $\Gamma$  acts transitively on the solutions of (1) modulo  $n$ , for every square free  $n \in \mathbb{N}$ . We give the precise formulation in Theorem 1.6 below.

Here we prove this conjecture for most primes. More particularly, we prove it for every  $p \equiv 1(4)$  outside the exceptional set from Theorem 1.1, and for density-1 of the primes  $p \equiv 3(4)$ :

---

<sup>1</sup>Sometimes the Markoff equation is written as  $x^2 + y^2 + z^2 = 3xyz$ . However, these two equations are equivalent in the sense that their integer solutions are related bijectively by  $(x, y, z) \longleftrightarrow (3x, 3y, 3z)$ . This bijection holds also for solutions in  $\mathbb{Z}/p\mathbb{Z}$  for every prime  $p \neq 3$ .

**Theorem 1.3.** *If  $p \equiv 1 (4)$  and  $Q_p$  is transitive, then  $Q_p$  is the full alternating or symmetric group on  $Y^*(p)$ .*

Namely,  $Q_p$  is the full alternating or symmetric group for all  $p \equiv 1 (4)$  outside the exceptional set from Theorem 1.1. In fact, our proof yields that for every  $p \equiv 1 (4)$ , the group  $\Gamma$  acts as the full alternating or symmetric group on the large component described in Theorem 1.1. In the case  $p \equiv 3 (4)$ , our proof is more involved and requires one further assumption:

**Theorem 1.4.** *Let  $p$  be a prime. Assume that:*

- $p \equiv 3 (4)$ .
- $Q_p$  is transitive.
- The order of  $\frac{3+\sqrt{5}}{2} \in \mathbb{F}_{p^2}$  is at least  $32\sqrt{p+1}$ .

*Then  $Q_p$  is the full alternating or symmetric group on  $Y^*(p)$ .*

As shown in Appendix A, the condition regarding the order of  $\frac{3+\sqrt{5}}{2}$  is satisfied for density-1 of the primes<sup>2</sup>, hence

**Corollary 1.5.** *For density-1 of all primes  $p \equiv 3 (4)$ , the group  $Q_p$  is the full alternating or symmetric group on  $Y^*(p)$ .*

## Composite Moduli

Let  $n$  be a positive integer which is square-free, so  $n = p_1 \cdots p_k$  where  $p_1, \dots, p_k$  are distinct primes. Let  $\mathbb{X}(n)$  denote the set of solutions to the Markoff equation (1) in  $\mathbb{Z}/n\mathbb{Z}$ . By the Chinese Remainder Theorem,  $\mathbb{X}(n) = \mathbb{X}(p_1) \times \dots \times \mathbb{X}(p_k)$ , and let  $X^*(n) = X^*(p_1) \times \dots \times X^*(p_k)$  be the set of solutions which are non-zero modulo any of the primes composing  $n$ . The action of  $\Gamma$  on  $\mathbb{X}(n)$  is the diagonal action on the  $\mathbb{X}(p_i)$ , and the subset  $X^*(n)$  is invariant under this action. Denote the corresponding permutation group  $\Gamma_n$ . Is the action on  $X^*(n)$  transitive? It turns out that this would follow from Conjecture 1.2 and indeed holds true for the cases of that conjecture we establish:

**Theorem 1.6.** *Let  $n = p_1 \cdots p_k$  be a product of distinct primes. If  $Q_{p_1}, \dots, Q_{p_k}$  are all alternating or symmetric, then  $\Gamma$  acts transitively on  $X^*(n)$ .*

*In particular, if conjecture 1.2 holds, then  $\Gamma$  acts transitively on  $X^*(n)$  for every square-free  $n$ .*

**Corollary 1.7.** *Let  $\mathcal{P}$  denote the set of primes that satisfy the assumptions of Theorem 1.3 or of Theorem 1.4. Then for every set of distinct primes  $p_1, \dots, p_k \in \mathcal{P}$ ,  $\Gamma$  acts transitively on  $X^*(p_1 \cdots p_k)$ .*

The second paper of Bourgain, Gamburd and Sarnak in the series announced in [BGS16] should also contain a result in the spirit of Corollary 1.7 for primes  $p \equiv 1 (4)$  for which  $\Gamma_p$  is transitive.

For  $n = p_1 \cdots p_k$  as above, we use the notation  $Y^*(n) = Y^*(p_1) \times \dots \times Y^*(p_k)$  for the set of blocks in  $X^*(n)$  and  $Q_n$  for the permutation group induced by the action of  $\Gamma$  on  $Y^*(n)$ . Note that these blocks are given by sign changes modulo every prime separately and are usually of size  $4^k$  each (if all primes are odd). It is quite straight-forward to prove that under the assumptions of Theorem 1.6,  $\Gamma$  acts transitively on  $Y^*(n)$ , using composition factors of  $Q_n$ . It requires some further argument to show that  $\Gamma$  acts transitively on the full set  $X^*(n)$ .

---

<sup>2</sup>A set of primes  $\mathcal{A}$  has *density 1* if  $\lim_{n \rightarrow \infty} \frac{|\mathcal{A} \cap \mathcal{P}_n|}{|\mathcal{P}_n|} = 1$ , where  $\mathcal{P}_n = \{1 < p \leq n \mid p \text{ is prime}\}$ . In fact, the set of primes for which  $\frac{3+\sqrt{5}}{2}$  has order at least  $32\sqrt{p+1}$  satisfies something slightly stronger than density 1 – see Appendix A.

*Remark 1.8* (Regarding the classification of finite simple groups). At this point we would like to remark on the dependence of our results on the Classification of Finite Simple Groups (CFSG). We use this theorem only in the proof of Theorem 1.4: we first give an elementary proof that for a prime  $p$  satisfying the assumptions in the theorem,  $Q_p$  is a primitive permutation group<sup>3</sup>, and then rely on (results depending on) the CFSG to deduce that  $Q_p$  is the full alternating or symmetric group. If we rely on Theorem 1.6 to deduce Corollary 1.7, the latter also becomes partly dependent on the CFSG. This can be avoided, however, and to this aim we also give a proof that  $\Gamma$  acts transitively on  $X^*(n)$  assuming only that  $Q_{p_1}, \dots, Q_{p_k}$  are primitive permutation groups, without using the CFSG (see Theorem 1.9 below). To sum up, the only results depending on the CFSG are Theorem 1.4, Corollary 1.5, and the part of Theorem 1.11 relating to primes  $p \equiv 3(4)$ . In contrast, Theorems 1.3 and 1.6 and Corollary 1.7 do not depend on the CFSG.

Indeed, the following result does not depend on the CFSG:

**Theorem 1.9.** *Let  $n = p_1 \cdots p_k$  be a product of distinct primes. If  $Q_{p_1}, \dots, Q_{p_k}$  are primitive permutation groups, then  $\Gamma$  acts transitively on  $X^*(n)$ .*

## $T_2$ -systems

Let  $G$  be a finitely generated group and  $F_r$  the free group on  $r$  generators. A normal subgroup  $N \trianglelefteq F_r$  is said to be  $G$ -defining if  $F_r/N \cong G$ . Denote by  $\Sigma_r(G)$  the set of  $G$ -defining normal subgroups in  $F_r$ . Consider the action of  $\text{Aut}(F_r)$  (in fact, of  $\text{Out}(F_r)$ ) on  $\Sigma_r(G)$ . The orbits of this action are called  $T_r$ -systems of  $G$ .

The following theorem is essentially due to Gilman:

**Theorem 1.10.** *[Gil77] Let  $G$  be a finite non-abelian simple group and  $r \geq 3$ . Then  $\text{Aut}(F_r)$  acts on at least one  $T_r$ -system of  $G$  as the alternating or symmetric group.*

Gilman also showed that for  $G = \text{PSL}(2, p)$  with  $p \geq 5$  prime, there is only one  $T_r$ -system for  $r \geq 3$ . Namely, he proved that  $\text{Aut}(F_r)$  acts transitively on  $\Sigma_r(G)$ . Theorem 1.10 says, of course, that the permutation group in this case is the alternating or symmetric group. For more details we refer the reader to the beautiful surveys [Pak01, Lub11].

When  $r = 2$ , the action of  $\text{Aut}(F_2)$  on  $\Sigma_2(G)$  is not transitive for any finite non-abelian simple group  $G$ . In fact, the number of  $T_2$ -systems tends to infinity as  $|G| \rightarrow \infty$  [GS09]. The main reason for this phenomenon is that, roughly, if  $\{a, b\}$  are a set of generators of  $F_2$ , and  $\varphi: F_2 \twoheadrightarrow G$  an epimorphism, then the conjugacy class of  $\varphi([a, b])$  is a well-defined invariant of the  $G$ -defining subgroup  $N = \ker \varphi$ , which is also invariant under  $\text{Aut}(F_2)$ . We elaborate more in Section 6.

Our result sheds more light on the case of  $T_2$ -systems for  $G = \text{PSL}(2, p)$ . If  $A, B \in \text{SL}(2, p)$  and we denote  $x = \text{tr}(A)$ ,  $y = \text{tr}(B)$  and  $z = \text{tr}(AB)$ , then

$$\text{tr}([A, B]) = x^2 + y^2 + z^2 - xyz - 2.$$

Moreover, if  $\langle A, B \rangle = \text{SL}(2, p)$  then the values  $\text{tr}(A)$ ,  $\text{tr}(B)$ ,  $\text{tr}(AB)$  uniquely determine  $N$ , the kernel of the epimorphism  $F_2 \twoheadrightarrow \text{PSL}(2, p)$  given by mapping  $a \mapsto A$ ,  $b \mapsto B$  (see Section 6). Thus, solutions to the Markoff equation (1) in  $\mathbb{Z}/p\mathbb{Z}$  correspond to elements in  $\Sigma_2(\text{PSL}(2, p))$  with associated trace  $-2$ . In this language, the main result of [BGS17] – Theorem 1.1 above – says that outside the exceptional set of primes, these elements form a single  $T_2$ -system. See [MW13] for an extensive survey of the connection between the Markoff equation (1) and  $T_2$ -systems of  $\text{PSL}(2, p)$ . Through this connection, Theorems (1.3) and (1.4) translate to a result in the spirit of Theorem 1.10:

---

<sup>3</sup>Recall that a permutation group  $G \leq \text{Sym}(m)$  is called primitive if it does not preserve any non-trivial block-structure.

**Theorem 1.11.** *Assume that the prime  $p$  satisfies the assumptions of Theorem 1.3 or of Theorem 1.4. Then  $\text{Aut}(\mathbb{F}_2)$  acts on the trace-(-2)  $T_2$ -system of  $\text{PSL}(2, p)$  as the full alternating or symmetric group.*

The paper is organized as follows. Section 2 gives some more notation and collects some results from [BGS17] we use here. In the short Section 3 and longer Section 4 we prove Theorem 1.3 for  $p \equiv 1(4)$  and Theorem 1.4 for  $p \equiv 3(4)$ , respectively. Section 5 is dedicated to proving the transitivity of  $\Gamma$  in certain composite moduli: first assuming the groups  $Q_p$  contain the alternating group (in Section 5.1), and then assuming only that  $Q_p$  is primitive (Section 5.2). In Section 6 we give some background on  $T$ -systems and prove Theorem 1.11. Finally, Appendix A, by Dan Carmon, shows that the assumption in Theorem 4 regarding the order of  $\frac{3+\sqrt{5}}{2} \in \mathbb{F}_{p^2}$  holds for most primes.

## Acknowledgments

We are in debt to Peter Sarnak for his encouragement, and for stimulating discussions, enlightening suggestions and clever advice. We would also like to thank Zeev Rudnick and Pär Kurlberg for beneficial comments, and to Dan Carmon for writing the useful Appendix A. We have benefited much from the mathematical open source community, and in particular from SageMath. Author D.P. was supported by the Rothschild fellowship, by the NSF under agreement No. DMS-1128155 and by the ISF grant 1071/16. Author D.C. was supported by the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement n° 320755.

## 2 Preliminaries

Before proving our main results, let us describe some further notation and collect further results from [BGS17] that we use below.

### Further notation

- We already introduced above the notation  $[x, y, z]$  for the block of the solution  $(x, y, z)$  in  $X^*(p)$ , so  $[x, y, z] \in Y^*(p)$ . We also use this notation for composite (square-free) modulo  $n$ : here  $[x, y, z]$  is the element (block) in  $Y^*(n)$  containing the solution  $(x, y, z)$ .
- Some elements in  $\Gamma$  are permutations of the three coordinates of solutions. We denote these elements by  $\tau_{(12)}$  for the permutation exchanging the first and second coordinates,  $\tau_{(123)}$  for the cyclic permutation and so on. By abuse of notation, we use the same notation for the corresponding elements in  $\Gamma$ ,  $\Gamma_p$ ,  $Q_p$ ,  $\Gamma_n$  and  $Q_n$ .
- The analysis in [BGS17], as well as in the current work, relies heavily on three “rotation” elements  $\text{rot}_1, \text{rot}_2, \text{rot}_3 \in \Gamma$ . They are defined by  $\text{rot}_i$

$$\text{rot}_j \stackrel{\text{def}}{=} R_{j+2} \circ \tau_{(j+1 \ j+2)}$$

(the indices are taken modulo 3). For example,  $(x, y, z) \xrightarrow{\text{rot}_1} (x, z, xz - y)$ . The rotation  $\text{rot}_j$  fixes the  $j$ -th coordinate and its action on  $X^*(p)$  and on  $Y^*(p)$  is fully understood – see Lemmas 2.2 and 2.3 below. Again, by abuse of notation we write  $\text{rot}_i$  for the rotation element in the different groups  $\Gamma$ ,  $\Gamma_p$ ,  $Q_p$ ,  $\Gamma_n$  and  $Q_n$ .

- Following [BGS17], we denote the “conic sections” by  $C_j(a)$ ,  $j = 1, 2, 3$ . These are  $C_j(a)$

defined as

$$C_j(a) = \{(x_1, x_2, x_3) \in X^*(p) \mid x_j = a\}.$$

When we write  $C_j(\pm a)$ , we mean the conic section in  $Y^*(p)$ :

$$C_j(\pm a)$$

$$C_j(\pm a) = \{[x_1, x_2, x_3] \in Y^*(p) \mid x_j = a\}.$$

- For every prime  $p$  we let  $i$  denote a square root of  $-1$  (in  $\mathbb{F}_p$  or in  $\mathbb{F}_{p^2}$ ). i
- For  $x \in \mathbb{Z}/p\mathbb{Z}$  we use the standard Legendre symbol  $\left(\frac{x}{p}\right)$  to denote the image of  $x$  under the character of order 2.  $\left(\frac{x}{p}\right)$
- The notation  $|x|$  is used to denote the order of the group element  $x \in G$  in the group  $G$ .

## Rotation elements

The action of  $\text{rot}_1$  on the conic section  $C_1(x) \subseteq X^*(p)$  is a linear map on the last two coordinates given by the matrix

$$\begin{pmatrix} 0 & 1 \\ -1 & x \end{pmatrix}. \quad (2)$$

The eigenvalues of this matrix are given by  $\frac{x \pm \sqrt{x^2 - 4}}{2}$ . This leads to the following definitions and lemmas from [BGS17]:

- Definition 2.1.**
- An element  $x \in \mathbb{F}_p$  is called *hyperbolic* if  $(x^2 - 4)$  is a square in  $\mathbb{F}_p^*$ . *hyperbolic*
  - An element  $x \in \mathbb{F}_p$  is called *elliptic* if  $(x^2 - 4)$  is a non-square in  $\mathbb{F}_p^*$ . *elliptic*
  - An element  $x \in \mathbb{F}_p$  is called *parabolic* if  $(x^2 - 4) = 0$  in  $\mathbb{F}_p$ , namely, if  $x = \pm 2$ . *parabolic*

Notice that this categorization of the elements is invariant under sign change  $x \mapsto -x$ . The following lemmas are based on Lemmas 3-5 of [BGS17] which describe the action of  $\text{rot}_i$  on  $X^*(p)$ . We adapt them below in order to describe the action of  $\text{rot}_i$  on  $Y^*(p)$  and add some further details, all follow easily from Section 2.1 in [BGS17]. We state the lemmas for  $C_1(\pm x)$ , but the same statements hold, evidently, for  $C_2(\pm x)$  and for  $C_3(\pm x)$ .

**Lemma 2.2.** [BGS17, Lemmas 3-5] *Let  $p \equiv 1 \pmod{4}$  be prime. Then,*

- $|Y^*(p)| = \frac{p(p+3)}{4}$ .
- $|C_1(\pm 2)| = p$ ; The permutation induced by  $\text{rot}_1$  on  $C_1(\pm 2)$  consists of a single  $p$ -cycle.
- There are  $\frac{p-1}{4}$  hyperbolic elements up to sign. For  $x$  hyperbolic,  $|C_1(\pm x)| = \frac{p-1}{2}$ . Let  $\omega^{\pm 1} \in \mathbb{F}_p$  be the eigenvalues of the matrix (2), so  $x = \omega + \omega^{-1}$ . The permutation induced by  $\text{rot}_1$  on  $C_1(\pm x)$  consists of  $\frac{p-1}{2d}$  cycles of length  $d$  each, where  $d = \frac{\max(|\omega|, |-\omega|)}{2}$  and  $|\omega|$  is the order of  $\omega$  in the multiplicative group  $\mathbb{F}_p^*$ . The solutions in  $C_1(x)$  have the form  $(x, \alpha + \beta, \alpha\omega + \beta\omega^{-1})$  for  $\alpha, \beta \in \mathbb{F}_p^*$  with  $\alpha\beta = \frac{x^2}{x^2 - 4}$ , and

$$(x, \alpha + \beta, \alpha\omega + \beta\omega^{-1}) \xrightarrow{\text{rot}_1} (x, \alpha\omega + \beta\omega^{-1}, \alpha\omega^2 + \beta\omega^{-2}). \quad (3)$$

- There are  $\frac{p-1}{4}$  elliptic elements up to sign. For  $x$  elliptic,  $|C_1(\pm x)| = \frac{p+1}{2}$ . Define  $\omega$  as for hyperbolic elements by  $x = \omega + \omega^{-1}$ , only now  $\omega \in \mathbb{F}_{p^2} \setminus \mathbb{F}_p$ . The permutation induced by  $\text{rot}_1$  on  $C_1(\pm x)$  consists of  $\frac{p+1}{2d}$  cycles of length  $d$  each, where  $d = \frac{\max(|\omega|, |-\omega|)}{2}$  and  $|\omega|$  is the order of  $\omega$  in the multiplicative group  $\mathbb{F}_{p^2}^*$ . Note that  $\omega^{p+1} = 1$ , so



type of $x$	# $x$ 's up to sign	$ C_1(\pm x) $		cycle-structure for $\text{rot}_1 _{C_1(\pm x)}$
parabolic	1	$p$	$x = \pm 2$	a single $p$ -cycle
hyperbolic (including 0)	$\frac{p-1}{4}$	$\frac{p-1}{2}$	$x = \omega + \omega^{-1},$ $\omega \in \mathbb{F}_p^* \setminus \{\pm 1\}$	For every $1 \neq d \mid \frac{p-1}{2}$ , there are $\left\lfloor \frac{\varphi(d)}{2} \right\rfloor$ hyperbolic $\pm x$ such that $\text{rot}_1 _{C_1(\pm x)}$ has $\frac{p-1}{2d}$ cycles of length $d$ each. (If $d$ is odd, $ w  \in \{d, 2d\}$ , if $d$ is even, $ w  = 2d$ .)
elliptic	$\frac{p-1}{4}$	$\frac{p+1}{2}$	$x = \omega + \omega^{-1}$ $\omega \in \mathbb{F}_{p^2}^* \setminus \{\pm 1\}$ $\omega^{p+1} = 1$	For every $1 \neq d \mid \frac{p+1}{2}$ , there are $\frac{\varphi(d)}{2}$ elliptic $\pm x$ such that $\text{rot}_1 _{C_1(\pm x)}$ has $\frac{p+1}{2d}$ cycles of length $d$ each. ( $ \omega  \in \{d, 2d\}$ )

Table 1: The structure of  $\text{rot}_1 \in Q_p$  when  $p \equiv 1 (4)$

$|\omega| \mid (p+1)$ . The solutions in  $C_1(x)$  have the form  $(x, A + A^p, A\omega + A^p\omega^{-1})$  with  $A \in \mathbb{F}_{p^2}^*$  and  $A^{p+1} = \frac{x^2}{x^2-4}$ , and

$$(x, A + A^p, A\omega + A^p\omega^{-1}) \xrightarrow{\text{rot}_1} (x, A\omega + A^p\omega^{-1}, A\omega^2 + A^p\omega^{-2}). \quad (4)$$

We sum up the content of Lemma 2.2 in Table 1.

Our results are somewhat weaker and the proof more involved when  $p \equiv 3 (4)$  because there are no solutions with the parabolic elements  $\pm 2$ :

**Lemma 2.3.** [BGS17, Lemmas 3-5] *Let  $p \equiv 3 (4)$  be prime. Then,*

- $|Y^*(p)| = \frac{p(p-3)}{4}$
- There are no solutions in  $Y^*(p)$  involving the parabolic elements  $\pm 2$ , nor the elliptic element 0.
- There are  $\frac{p-3}{4}$  hyperbolic elements up to sign. For  $x$  hyperbolic, the size and structure of  $C_1(\pm x)$  and the action of  $\text{rot}_1$  on  $C_1(\pm x)$  have the same properties as for  $x$  hyperbolic when  $p \equiv 1 (4)$  (see Lemma 2.2).
- There are  $\frac{p-3}{4}$  non-zero elliptic elements up to sign. For  $x$  elliptic, the size and structure of  $C_1(\pm x)$  and the action of  $\text{rot}_1$  on  $C_1(\pm x)$  have the same properties as for  $x$  elliptic when  $p \equiv 1 (4)$  (see Lemma 2.2).

We sum up the content of Lemma 2.3 in Table 2.

For  $x \in \mathbb{F}_p$ , denote by  $d_p(\pm x)$  the order of  $\text{rot}_1 \in Q_p$  in its action on  $C_1(\pm x)$ . Namely,  $d_p(\pm x)$  the solutions with first coordinate  $\pm x$  in  $Y^*(p)$  belong to cycles of length  $d_p(\pm x)$ .

### 3 Alternating Group for $p \equiv 1 (4)$

This section contains the proof of Theorem 1.3, which states that if  $p \equiv 1 (4)$  and  $Q_p$  is transitive, then  $Q_p$  contains the entire alternating group  $\text{Alt}(Y^*(p))$ . As mentioned above, the existence of parabolic elements when  $p \equiv 1 (4)$  allows a rather short argument in this case.

We use the following classical theorem of Jordan:

type of $x$	# $x$ 's up to sign	$ C_1(\pm x) $	eigenvalues of $\text{rot}_1$	cycle-structure of $\text{rot}_1 _{C_1(\pm x)}$
hyperbolic $\left(\frac{x^2-4}{p}\right) = 1$	$\frac{p-3}{4}$	$\frac{p-1}{2}$	$\omega \in \mathbb{F}_p^* \setminus \{\pm 1\}$ $x = \omega + \omega^{-1}$	For every $1 \neq d \mid \frac{p-1}{2}$ , there are $\frac{\varphi(d)}{2}$ hyperbolic $\pm x$ such that $\text{rot}_1 _{C_1(\pm x)}$ has $\frac{p-1}{2d}$ cycles of length $d$ each. ( $ w  \in \{d, 2d\}$ )
elliptic (exc. 0): $x \neq 0$ & $\left(\frac{x^2-4}{p}\right) = -1$	$\frac{p-3}{4}$	$\frac{p+1}{2}$	$\omega \in \mathbb{F}_{p^2}^* \setminus \{\pm 1, \pm i\}$ $x = \omega + \omega^{-1}$ $\omega^{p+1} = 1$	For every $3 \leq d \mid \frac{p+1}{2}$ , there are $\frac{\varphi(d)}{2}$ elliptic $\pm x$ such that $\text{rot}_1 _{C_1(\pm x)}$ has $\frac{p+1}{2d}$ cycles of length $d$ each. (If $d$ is odd, $ \omega  \in \{d, 2d\}$ , if $d$ is even, $ \omega  = 2d$ .)

Table 2: The structure of  $\text{rot}_1 \in Q_p$  when  $p \equiv 3(4)$

**Theorem 3.1** (Jordan [DM96, Thm 3.3E]). *Let  $G \leq \text{Sym}(n)$  be a primitive permutation group containing a cycle of prime length  $p \leq n - 3$ . Then  $G \geq \text{Alt}(n)$ .*

*Proof of Theorem 1.3.* Assume  $p \equiv 1(4)$ , and let  $\text{rot}_1 \in Q_p$  be the rotation element defined on Page 5. This element has one  $p$ -cycle, while all its other cycles have length coprime to  $p$  (see Table 1). Thus its power  $\sigma = \text{rot}_1^{|\text{rot}_1|/p} \in Q_p$  is a  $p$ -cycle. As  $|Y^*(p)| = \frac{p(p+3)}{4} \geq p + 3$ , it is now sufficient to show, by Jordan's Theorem (Theorem 3.1 above), that  $Q_p$  is primitive in  $\text{Sym}(Y^*(p))$ .

The group  $Q_p$  is transitive by assumption, so we only need to show it preserves no non-trivial block structure. Assume there is a block structure  $\{B_1, \dots, B_m\}$  preserved by  $Q_p$ . So  $\bigcup B_i = Y^*(p)$  and  $B_i \cap B_j = \emptyset$  for  $i \neq j$ , and for every  $g \in Q_p$  and every  $i$ ,  $g(B_i) = B_j$  for some  $j$ .

Consider  $C_1(\pm 2) \subset Y^*(p)$ , the  $p$  elements contained in the cycle of size  $p$  in  $\sigma$ . The set  $C_1(\pm 2)$  must be contained in a block, for otherwise it has to be the union of several equally-sized blocks, but  $p$  is prime. Say  $C_1(\pm 2) \subseteq B_1$ . So  $B_1$  contains all solutions with  $\pm 2$  in the first coordinate. In particular, it contains  $[2, 2, 2 + 2i]$  and  $[2, 2 + 2i, 2]$ . But the same argument with  $\text{rot}_2$  and  $\text{rot}_3$  shows that  $B_1$  contains all solutions with  $\pm 2$  in any coordinate. So  $B_1$  is invariant not only under  $\text{rot}_1$  but also under any permutation of coordinates, and therefore invariant under the action of the whole group  $Q_p$ . By the transitivity of  $Q_p$ ,  $B_1 = Y^*(p)$ .  $\square$

*Remark 3.2.* The proof of Theorem 1.1 in [BGS17] shows that for every prime  $p$ , the large component of  $X^*(p)$  contains all solutions with parabolic ( $\pm 2$ ) coordinates. Thus, our proof of Theorem 1.3 applies to the general case: the group  $\Gamma$  acts on the large component of  $Y^*(p)$  as the alternating or symmetric group.

## 4 Alternating Group for $p \equiv 3(4)$

In the case where  $p \equiv 3(4)$ , there are no parabolic elements, and we establish the primitivity of  $Q_p$  for density-1 of these primes rather than for all those outside the exceptional set from Theorem 1.1. We also rely on much deeper theorems, involving the classification of finite simple groups (CFSG), to conclude that whenever  $Q_p$  is primitive, it contains  $\text{Alt}(Y^*(p))$ . Throughout this section, we assume that  $p \equiv 3(4)$ .



#### 4.1 Deducing Alternating group from primitivity

First, we show how to deduce that  $Q_p \geq \text{Alt}(Y^*(p))$  whenever  $Q_p$  is primitive. Throughout this section we denote the symmetric group  $\text{Sym}(n)$  by  $S_n$  and  $\text{Alt}(n)$  by  $A_n$ . Here we use the following result of Guralnick and Magaard, classifying primitive subgroups of  $S_n$  containing an element with at least  $n/2$  fixed points. This theorem relies heavily on the CFSG. We adjust the statement of the theorem to our needs – the original statement in [GM98] is more detailed. In the statement we use the notation  $\text{Soc}(G)$  for the socle of the group  $G$  (see Section 5.2 for details), and the standard notation  $G_1 \wr G_2$  for the wreath product of two groups.

**Theorem 4.1** ([GM98, Theorem 1]). *Let  $G \leq S_n$  be a primitive group, and let  $x \in G$  have at least  $n/2$  fixed points. Then one of the following holds:*

1.  $G = \text{Aff}(2, k)$  is the affine group acting on  $\mathbb{F}_2^k$  and  $x$  is a transvection<sup>4</sup> and is, in particular, an involution. In this case  $x$  has exactly  $n/2$  fixed points.
2. There are  $r \geq 1$ ,  $m \geq 5$  and  $1 \leq k \leq m/4$  such that  $n = \binom{m}{k}^r$ , the group  $S_m$  acts on the set  $\Delta$  of  $k$ -subsets of  $\{1, \dots, m\}$  in the natural way,  $G \leq S_m \wr S_r$  acts on  $\Delta^r$  and  $\text{Soc}(G) = A_m^r$ .
3. For some  $r \geq 1$ ,  $n = 6^r$ , the group  $S_6$  acts on  $\Delta = \{1, \dots, 6\}$  by applying an outer automorphism<sup>5</sup>,  $G \leq S_6 \wr S_r$  acts on  $\Delta^r$  and  $\text{Soc}(G) = A_6^r$ .
4. The group  $G$  is some variant of an orthogonal group over the field of two elements acting on some collection of 1-spaces or hyperplanes, and the element  $x$  is an involution.

The following lemma helps us rule out Case 2 of the above theorem with  $r = 1$ .

**Lemma 4.2.** *Consider the embedding  $\iota: S_m \hookrightarrow S_n$  given by the natural action of the symmetric group  $S_m$  on the set  $\Delta$  of  $n = \binom{m}{k}$   $k$ -subsets of  $m$ , for some  $2 \leq k \leq \frac{m}{4}$ . If, for some  $\pi \in S_m$ , the image  $\iota(\pi)$  has a cycle of size divisible by  $q$  and a cycle of size divisible by  $s$  for some distinct primes  $q$  and  $s$ , then  $\iota(\pi)$  also has a cycle of size divisible by  $qs$ .*

*Proof.* Assume that  $\{a_1, \dots, a_k\} \in \Delta$  belongs to a cycle  $\alpha$  of length divisible by  $q$  in  $\iota(\pi)$ . Assume that in  $\pi$ , the elements  $a_1, \dots, a_k$  belong to  $t$  distinct cycles: the elements  $a_1, \dots, a_{\ell_1}$  belong to the cycle  $\sigma_1$ , the elements  $a_{\ell_1+1}, \dots, a_{\ell_2}$  belong to the cycle  $\sigma_2$ , and so on. Let  $o_1$  be the smallest power of  $\sigma_1$  that maps  $\{a_1, \dots, a_{\ell_1}\}$  to itself. Define  $o_2, \dots, o_t$  analogously. Then,  $q \mid \text{lcm}(o_1, \dots, o_t)$ . In particular,  $q \mid o_i$  for some  $i$ , and so  $q \mid |\sigma_i|$ . Without loss of generality, assume  $q \mid |\sigma_1|$ , so that  $a_1$  belongs to a cycle  $\sigma = \sigma_1$  of  $\pi$  of size divisible by  $q$ . Likewise, assume that  $b_1$  belongs to a cycle  $\tau$  of  $\pi$  of size divisible by  $s$ .

If  $\tau$  and  $\sigma$  are the same cycle, we are done, so assume otherwise. Denote  $A = \{1, \dots, m\} \setminus (\sigma \cup \tau)$ . If  $|A| \geq k - 2$ , then a  $k$ -subset containing  $a_1, b_1$  and  $k - 2$  elements from  $A$  belongs to a cycle of  $\iota(\pi)$  of size divisible by  $qs$ . If  $|A| < k - 2$ , then, as  $k \leq \frac{m}{4}$ , at least one of  $\sigma$  or  $\tau$  has more than  $k$  element. Assume without loss of generality it is  $\sigma$ . Consider the  $k$ -subset  $\{b_1, a_1, \pi(a_1), \pi^2(a_1), \dots, \pi^{k-2}(a_1)\}$ . This subset belongs to a cycle of  $\iota(\pi)$  of size  $\text{lcm}(|\tau|, |\sigma|)$ , which, in particular, is a multiple of  $qs$ .  $\square$

**Proposition 4.3.** *Let  $p \equiv 3(4)$  be prime. If  $Q_p$  is primitive, then  $Q_p \geq \text{Alt}(Y^*(p))$ .*

<sup>4</sup>To be sure,  $x$  is a transvection when  $\text{Aff}(2, k)$  is embedded in  $\text{GL}(2, k + 1)$  as the matrices with bottom row  $(0, \dots, 0, 1)$ .

<sup>5</sup>Namely, for some fixed  $\varphi \in \text{Aut}(S_6) \setminus \text{Inn}(S_6)$ , the permutation  $\sigma \in S_6$  acts on  $\Delta$  by  $\sigma.i = \varphi(\sigma)(i)$ .

*Proof.* Consider  $\text{rot}_1 \in Q_p$ . Among the  $\frac{p(p-3)}{4}$  elements in  $Y^*(p)$ ,  $\frac{(p-1)(p-3)}{8}$  belong to cycles of length at least 3 and dividing  $\frac{p-1}{2}$ , and  $\frac{(p+1)(p-3)}{8}$  belong to cycles of length at least 3 and dividing  $\frac{p+1}{2}$  (see Table 8). Since  $\gcd\left(\frac{p-1}{2}, \frac{p+1}{2}\right) = 1$ , the permutation  $\sigma = \text{rot}_1^{(p+1)/2}$  fixes exactly  $\frac{(p+1)(p-3)}{8} > \frac{|Y^*(p)|}{2}$  elements of  $Y^*(p)$ . Thus  $Q_p$  satisfies the assumptions in Theorem 4.1. We can now rule out all options except for  $Q_p = \text{Alt}(Y^*(p))$  or  $Q_p = \text{Sym}(Y^*(p))$ .

Cases 1 and 4 are immediately ruled out because the permutation  $\sigma \in Q_p$  is not an involution. Case 2 with  $r \geq 2$  and Case 3 are immediately ruled out because  $|Y^*(p)| = \frac{p(p-3)}{4}$  is not a proper power nor equal to six. It remains to consider Case 2 with  $r = 1$ .

Let  $q$  be some prime factor of  $\frac{p-1}{2}$ , and let  $s$  be some prime factor of  $\frac{p+1}{2}$ . By Table 2,  $\text{rot}_1$  contains cycles of size divisible by  $q$  (indeed, even of size  $q$  exactly), and of size divisible by  $s$ . However, it does not contain any cycle of size divisible by  $qs$ . Using Lemma 4.2, this rules out Case 2 from Theorem 4.1 with  $r = 1$  and  $k \geq 2$ . The remaining case, that of Case 2 with  $r = k = 1$ , is precisely the case that the group in question is either  $A_n$  or  $S_n$ .  $\square$

## 4.2 Primitivity for density-1 of the primes

We use the symmetric solution  $[3, 3, 3] \in Y^*(p)$  to derive primitivity. Relying on strong results of Ford [For08], Dan Carmon proves in Appendix A that for density-1 of the primes, the cycle in  $\text{rot}_1 \in Q_p$  containing this solution is fairly long (roughly, longer than  $\sqrt{p}$ ). We show below that in this case, if  $Q_p$  is transitive, it is also primitive.

More concretely, let  $\omega = \frac{3+\sqrt{5}}{2} \in \mathbb{F}_{p^2}$  (there are two such elements, pick one of them arbitrarily). Then  $3 = \omega + \omega^{-1}$ . Recall that  $d_p(\pm 3)$  denotes the length of the cycles of  $\text{rot}_1 \in Q_p$  containing elements of  $C_1(\pm 3)$ . By Lemma 2.3 and Table 2,  $d_p(\pm 3)$  is either  $|\omega|$  or  $\frac{|\omega|}{2}$ , where  $|\omega|$  is the order of  $\omega$  in the multiplicative group  $\mathbb{F}_{p^2}^*$ . From Proposition A.1 in Appendix A, we deduce:

**Proposition 4.4.** *For density-1 of all primes, the element  $\omega = \frac{3+\sqrt{5}}{2} \in \mathbb{F}_{p^2}$  has order at least  $32\sqrt{p+1}$  in the multiplicative group  $\mathbb{F}_{p^2}^*$ , in which case  $d_p(\pm 3) \geq 16\sqrt{p+1}$ .*

Combining Theorem 1.1 with Proposition 4.4 shows why the assumptions in Theorem 1.4 hold for density-1 of all primes  $p \equiv 3(4)$ . Thus, Corollary 1.5 follows from Theorem 1.4. It remains to show that for  $p \equiv 3(4)$ , if  $Q_p$  is transitive and  $d_p(\pm 3) \geq 16\sqrt{p+1}$ , then  $Q_p$  is primitive.

The following lemmas and proposition are the main building blocks in the proof.

**Lemma 4.5.** *The subgroup  $H = \langle \text{rot}_1, \text{rot}_2, \text{rot}_3 \rangle \leq \Gamma$  has index at most 2 in  $\Gamma$ .*

*Proof.* By definition,  $\Gamma$  is generated by the three Vieta involutions and permutations of coordinates. Since  $R_3 = \text{rot}_1 \cdot \tau_{(2\ 3)}$  and likewise for  $R_1$  and  $R_2$ , since  $\tau_{(1\ 3\ 2)} = \text{rot}_3 \cdot \text{rot}_1$  and since  $S_3 = \langle (12), (132) \rangle$ , we obtain that  $\Gamma = \langle \text{rot}_1, \text{rot}_2, \text{rot}_3, \tau_{(1\ 2)} \rangle = \langle H, \tau_{(1\ 2)} \rangle$ . It is easy to check that  $\tau_{(1\ 2)} \text{rot}_j \tau_{(1\ 2)} \in H$  for  $j = 1, 2, 3$ , so  $H \trianglelefteq \Gamma$  and  $\Gamma = H \cdot \langle \tau_{(1\ 2)} \rangle$ . This finishes the proof.  $\square$

We say that some coordinate  $j \in \{1, 2, 3\}$  is *homogeneous* in some block  $B \subseteq Y^*(p)$  if the  $j$ -th coordinate of every solution in  $B$  is of the same type (either all hyperbolic or all elliptic).

**Lemma 4.6.** *Let  $p \equiv 3(4)$ . Assume that  $Q_p$  acts transitively<sup>6</sup> on  $Y^*(p)$ , and let  $B \subsetneq Y^*(p)$*

<sup>6</sup>We assume transitivity only for simplicity. This assumption of Lemma 4.6 can be replaced by assuming that  $Q_p.B = \{\sigma(b) \mid \sigma \in Q_p, b \in B\}$  contains the solution  $[3, 3, 3]$  and  $B \subsetneq Q_p.B$ , with the exact same proof. It could also be replaced by assuming that  $Q_p.B$  contains some solution with some coordinate being hyperbolic of maximal order: if  $x = \omega + \omega^{-1}$  is hyperbolic with  $\omega$  of order  $p-1$ , then the proof could work using the pair of solutions  $\left[x, \frac{x^2}{\sqrt{x^2-4}}, \frac{2x}{\sqrt{x^2-4}}\right]$  and  $\left[x, \frac{2x}{\sqrt{x^2-4}}, \frac{x^2}{\sqrt{x^2-4}}\right]$ , which are connected by both  $\tau_{(1\ 2)}$  and  $\text{rot}_1$ .

be a proper block of the action of  $Q_p$  on  $Y^*(p)$ . Then at least two of the coordinates  $\{1, 2, 3\}$  are homogeneous in  $B$ .

*Proof.* Assume that some coordinate, say  $j = 1$ , is not homogeneous in  $B$ . We need to show that the second and third coordinates are homogeneous. The element  $\text{rot}_1^{(p-1)/2}$  fixes every solution with first coordinate hyperbolic, while  $\text{rot}_1^{(p+1)/2}$  fixes every solution with first coordinate elliptic. Hence  $B$  is invariant under both elements, and thus by  $\text{rot}_1$ .

By the same argument, if all three coordinates are not homogeneous,  $B$  is invariant under  $H_p = \langle \text{rot}_1, \text{rot}_2, \text{rot}_3 \rangle \leq Q_p$ . By Lemma 4.5,  $[Q_p : H_p] \leq 2$ , and transitivity implies there are at most two blocks in the action:  $B$  and  $B' = \gamma(B)$  for some  $\gamma \in Q_p$ . But the block containing  $[3, 3, 3]$  is also invariant under  $\tau_{(1\ 2)}$ , hence is invariant under the whole of  $Q_p$  - a contradiction.

Thus at least one coordinate - the second or the third - is homogeneous. Notice that  $\text{rot}_1$ , which stabilizes  $B$ , moves the third coordinate of the solutions to the second. Hence both the second and third coordinates must be homogeneous.  $\square$

*Remark 4.7.* In fact, the proof of the last lemma yields something slightly stronger. Denote the type of a solution in  $Y^*(p)$  by some triple in  $\{h, e\}^3$ , depending on whether every coordinate is hyperbolic or elliptic. Then, every block  $B$  as above contains either only solutions of the same type (homogeneous in all coordinates), or only solutions of exactly two types: one type is  $(h, h, h)$  or  $(e, e, e)$ , and the other differs from the first type in one coordinate (the sole non-homogeneous coordinate).

The most technical ingredient of the proof of primitivity is the following:

**Proposition 4.8.** *Assume that  $Q_p$  is transitive and that  $x \in \mathbb{F}_p \setminus \{0, \pm 2\}$  satisfies  $d_p(\pm x) \geq 16\sqrt{p+1}$ . Then, for every  $j \in \{1, 2, 3\}$ , every proper block  $B \subsetneq Y^*(p)$  of the action of  $Q_p$  on  $Y^*(p)$  contains at most one solution with  $j$ -th coordinate  $\pm x$ .*

The idea of the proof of this proposition is the following: assume there are two solutions in the block  $B$  with first coordinate  $\pm x$ . Say these are  $[x, y_0, y_1]$  and  $[x, z_0, z_1]$ . Then for every  $1 \leq m$ , the block  $\text{rot}_1^m(B)$  contains the solutions  $[x, y_m, y_{m+1}]$  and  $[x, z_m, z_{m+1}]$  with  $y_{m+1} = xy_m - y_{m-1}$  and  $z_{m+1} = xz_m - z_{m-1}$ . By Lemma 4.6, at least one of the two coordinates 2, 3 in every block is homogeneous, meaning that for every  $m$ , either  $y_m$  and  $z_m$  have the same type ( $h$  or  $e$ ), or  $y_{m+1}$  and  $z_{m+1}$  have the same type. Using classical results in number theory, we show such “high correlation” between two cycles of  $\text{rot}_1$  is impossible whenever these cycles are long enough.

We defer the details of the proof to Section 4.3, and assuming it, finish the proof of Theorem 1.4. We need the following corollary showing that elements of high order in the sense of Proposition 4.8 appear in the same block and the same coordinate only with other elements of the same type and the same order:

**Corollary 4.9.** *Assume that  $Q_p$  is transitive and that  $x \in \mathbb{F}_p \setminus \{0, \pm 2\}$  satisfies  $d_p(\pm x) \geq 16\sqrt{p+1}$ . If  $B \subsetneq Y^*(p)$  is a proper block of the action of  $Q_p$  on  $Y^*(p)$  containing some solution with first coordinate  $\pm x$ , and another solution with first coordinate  $\pm x'$ , then  $d_p(\pm x) = d_p(\pm x')$ . In particular,  $x$  and  $x'$  are of the same type.*

*Proof.* Note that  $\text{rot}_1^{d_p(\pm x)}(B) = B$ . By Proposition 4.8,  $\text{rot}_1^m(B) \neq B$  for  $1 \leq m < d_p(\pm x)$ . Hence,  $d_p(\pm x')$  is some multiple of  $d_p(\pm x)$ . In particular, the assumption of Proposition 4.8 holds for  $x'$ , and by symmetry,  $d_p(\pm x)$  is a multiple of  $d_p(\pm x')$ . Hence  $d_p(\pm x') = d_p(\pm x)$ .  $\square$

*Proof of Theorem 1.4 assuming Proposition 4.8.* Assume that  $Q_p$  is transitive and  $\frac{3+\sqrt{5}}{2} \in \mathbb{F}_{p^2}^*$  has order at least  $32\sqrt{p+1}$ . We need to show that  $Q_p$  is primitive. Assume that  $[a, b, c]$  and

$[3, 3, 3]$  are two distinct solutions lying in the same block  $B$  of the action of  $Q_p$  on  $Y^*(p)$ . By Lemma 2.3,  $d_p(\pm 3) \geq 16\sqrt{p+1}$ , and by Corollary 4.9,  $d_p(\pm a) = d_p(\pm b) = d_p(\pm c) = d_p(\pm 3)$ . As  $[3, 3, 3]$  is the only solution of the form  $[x, x, x]$  or  $[x, x, -x]$ , we can assume without loss of generality that  $\{\pm b\} \neq \{\pm c\}$ . Since  $\tau_{(2\ 3)}$  stabilizes  $[3, 3, 3]$ , we have  $\tau_{(2\ 3)}(B) = B$ , so the two distinct solutions  $[a, b, c]$  and  $[a, c, b]$  both belong to  $B$ . This contradicts Proposition 4.8:  $d_p(\pm a) = d_p(\pm 3)$  is large and thus  $a$  cannot appear twice in the same coordinate in the same block.  $\square$

### 4.3 Ruling out correlation between two long $\text{rot}_1$ -cycles with same first coordinate

We now prove Proposition 4.8 stating the lack of correlation between two long enough cycles of  $\text{rot}_j$  with the same  $j$ -th coordinate (including the case of two different offsets of the same cycle). We use the following classical number-theoretic result:

**Theorem 4.10** (Weil [Sch76, Theorem II.2C']). *Let  $f(x) \in \mathbb{F}_p[x]$  be a polynomial with  $m$  distinct roots in  $\overline{\mathbb{F}_p}$  which is not a square in  $\overline{\mathbb{F}_p}[x]$ . Then*

$$\left| \sum_{s \in \mathbb{F}_p} \left( \frac{f(s)}{p} \right) \right| \leq (m-1) \sqrt{p}.$$

#### 4.3.1 Hyperbolic elements of high order

We state the following proposition for the first coordinate, but it holds, evidently, for every coordinate  $j = 1, 2, 3$ .

**Proposition 4.11.** *Assume  $Q_p$  is transitive and let  $x \in \mathbb{F}_p$  be hyperbolic with  $d = d_p(\pm x) \geq 16\sqrt{p-1}$ . If  $B \subsetneq Y^*(p)$  is a proper block of the action of  $Q_p$  on  $Y^*(p)$ , then  $B$  contains at most one solution with first coordinate  $\pm x$ .*

*Proof.* Assume that  $[x, y_0, y_1]$  and  $[x, z_0, z_1]$  belong to  $B$ . By Lemma 2.3,  $x = \omega + \omega^{-1}$  with  $\omega \in \mathbb{F}_p^*$  and we can assume  $|\omega| = 2d \geq 32\sqrt{p-1}$ : otherwise, replace  $x$  with  $-x$  and  $\omega$  with  $-\omega$ . Write  $y_0 = \alpha + \beta$  with  $\alpha, \beta \in \mathbb{F}_p^*$  so that  $\alpha\beta = \frac{x^2}{x^2-4}$  and  $y_1 = \alpha\omega + \beta\omega^{-1}$  (see Lemma 2.3). The cycle of  $\text{rot}_1$  containing  $[x, y_0, y_1]$  is

$$[x, y_0, y_1] = [x, y_d, y_{d+1}], [x, y_1, y_2], \dots, [x, y_{d-2}, y_{d-1}], [x, y_{d-1}, y_d]$$

with

$$y_j = \alpha\omega^j + \beta\omega^{-j}.$$

The set  $\{\omega^j\}_{0 \leq j \leq 2d-1}$  is the same as the set  $\{s^m\}_{s \in \mathbb{F}_p^*}$  where  $m = \frac{p-1}{2d}$  (with every element in  $\{\omega^j\}$  covered by  $\frac{p-1}{2d}$  different values of  $s$ ). So as sets,

$$\{y_0, \dots, y_{2d-1}\} = \{\alpha\omega^j + \beta\omega^{-j}\}_{0 \leq j \leq 2d-1} = \left\{ f_{\alpha, \beta}(s) \stackrel{\text{def}}{=} \alpha s^m + \beta s^{-m} \right\}_{s \in \mathbb{F}_p^*}.$$

The same holds for the cycle of  $\text{rot}_1$  containing  $[x, z_0, z_1]$  with  $\gamma, \delta \in \mathbb{F}_p^*$  in the role of  $\alpha, \beta$ , so that  $z_j = \gamma\omega^j + \delta\omega^{-j}$ . We may assume that  $\gamma \neq \pm\alpha$ , for otherwise  $[x, y_0, y_1] = [x, z_0, z_1]$ . Moreover, if  $s^m = \omega^j$  then  $f_{\alpha, \beta}(s) = y_j$  and  $f_{\gamma, \delta}(s) = z_j$ .

Notice that  $y_j$  and  $z_j$  have different type (one hyperbolic and the other elliptic) if and only if

$$\left( \frac{(y_j^2 - 4)(z_j^2 - 4)}{p} \right) = -1. \quad (5)$$

Since  $[x, y_j, y_{j+1}]$  and  $[x, z_j, z_{j+1}]$  both belong to the block  $\text{rot}_1^j(B)$ , we derive from Lemma 4.6 that (5) cannot hold for two consecutive values of  $j$ . In the parametrization given by  $s \in \mathbb{F}_p^*$ , this means that

$$\left( \frac{(f_{\alpha,\beta}(s)^2 - 4)(f_{\gamma,\delta}(s)^2 - 4)}{p} \right) = \left( \frac{(f_{\alpha\omega,\beta\omega^{-1}}(s)^2 - 4)(f_{\gamma\omega,\delta\omega^{-1}}(s)^2 - 4)}{p} \right) = -1 \quad (6)$$

cannot hold for any  $s \in \mathbb{F}_p^*$ .

Write

$$g_{\alpha,\beta}(s) \stackrel{\text{def}}{=} (s^m)^2 (f_{\alpha,\beta}(s)^2 - 4) = [(\alpha s^{2m} + \beta)^2 - 4s^{2m}] \in \mathbb{F}_p[s],$$

and  $k_1(s) = g_{\alpha,\beta}(s)g_{\gamma,\delta}(s)$  and  $k_2(s) = g_{\alpha\omega,\beta\omega^{-1}}(s)g_{\gamma\omega,\delta\omega^{-1}}(s)$ . Now (6) is equivalent to

$$\left( \frac{k_1(s)}{p} \right) = \left( \frac{k_2(s)}{p} \right) = -1. \quad (7)$$

For  $\ell \in \{\pm 1\}^2$ , denote by  $N_\ell$  the number of  $s \in \mathbb{F}_p$  for which  $\left( \frac{k_1(s)}{p} \right) = \ell_1$  and  $\left( \frac{k_2(s)}{p} \right) = \ell_2$ . Our goal is to show that  $N_{(-1,-1)} \geq 2$ , whence (7) has some solution  $s \neq 0$ , yielding a contradiction. For this goal, for every  $B \subseteq \{1, 2\}$ , let  $M_B$  denote the difference between the number of  $s \in \mathbb{F}_p$  for which  $\left( \frac{\prod_{j \in B} k_j(s)}{p} \right)$  has the same sign as for a solution of (7) and the number of these with opposite sign, namely

$$M_B \stackrel{\text{def}}{=} (-1)^{|B|} \sum_{s \in \mathbb{F}_p} \left( \frac{\prod_{j \in B} k_j(s)}{p} \right). \quad (8)$$

Consider the sum  $\sum_{B \subseteq \{1, 2\}} M_B$ . Every  $s \in \mathbb{F}_p$  which is a solution to (7) contributes 4 to this sum. Note that  $g_{\alpha,\beta}$  has no roots in  $\mathbb{F}_p$ : certainly  $g_{\alpha,\beta}(0) = \beta^2 \neq 0$ , and if  $0 \neq s \in \mathbb{F}_p$  and  $g_{\alpha,\beta}(s) = 0$  then  $f_{\alpha,\beta}(s) = \pm 2$  is  $y_j$  for some  $j$ , but there are no solution in  $X^*(p)$  containing  $\pm 2$  when  $p \equiv 3(4)$ . Thus, every  $s \in \mathbb{F}_p$  which is a non-solution to (7), contributes exactly 0 to this sum. We conclude that

$$N_{(-1,-1)} = \frac{1}{4} \left( \sum_{B \subseteq \{1, 2\}} M_B \right). \quad (9)$$

We use Theorem 4.10 to estimate the  $M_B$ 's. First, we show that none of  $k_1, k_2$  and  $k_1 k_2$  are squares in  $\overline{\mathbb{F}_p}[x]$ . The roots of

$$g_{\alpha,\beta}(s) = (\alpha s^{2m} + \beta + 2s^m)(\alpha s^{2m} + \beta - 2s^m)$$

satisfy

$$s^m = \frac{\pm 2 \pm \sqrt{4 - 4\alpha\beta}}{2\alpha} = \frac{\pm 1 \pm \sqrt{1 - \frac{x^2}{x^2 - 4}}}{\alpha} = \frac{\pm 1 \pm \sqrt{\frac{-4}{x^2 - 4}}}{\alpha}.$$

As  $x$  is hyperbolic and  $p \equiv 3(4)$ , we have that  $\frac{-4}{x^2 - 4}$  is not a square in  $\mathbb{F}_p$ , so 1 and  $\sqrt{\frac{-4}{x^2 - 4}}$  are linearly independent over  $\mathbb{F}_p$ , and  $\frac{\pm 1 \pm \sqrt{\frac{-4}{x^2 - 4}}}{\alpha}$  are four distinct values for  $S^m$ , different from zero. Moreover, the polynomial  $s^m - \xi$  is separable for  $0 \neq \xi \in \mathbb{F}_{p^2}$  because  $m = \frac{p-1}{2d} < p$ . So  $g_{\alpha,\beta}(s)$ , which is of degree  $4m$ , has  $4m$  distinct roots in  $\overline{\mathbb{F}_p}$ , and in particular is not a square in  $\overline{\mathbb{F}_p}[x]$ .

This analysis shows that  $g_{\alpha,\beta}$  and  $g_{\gamma,\delta}$  have a common root if and only if  $\alpha = \pm\gamma$ . Since  $\alpha \neq \pm\gamma$  by assumption,  $k_1 = g_{\alpha,\beta}g_{\gamma,\delta}$  and  $k_2 = g_{\alpha\omega,\beta\omega^{-1}}g_{\gamma\omega,\delta\omega^{-1}}$  are both separable of degree  $8m$ . Finally,  $k_1k_2$ , of degree  $16m$ , is also not a square in  $\mathbb{F}_p[x]$ : for  $\alpha \neq \pm\alpha\omega$  and if  $\alpha = \pm\gamma\omega$  then  $\alpha\omega \neq \pm\gamma$ .

Of course,  $M_\emptyset = p$ . Theorem 4.10 yields that  $|M_{\{1\}}|, |M_{\{2\}}| \leq (8m-1)\sqrt{p}$  and  $|M_{\{1,2\}}| \leq (16m-1)\sqrt{p}$ . From (9) we now obtain

$$\begin{aligned} N_{(-1,-1)} &\geq \frac{1}{4} [p - 2(8m-1)\sqrt{p} - (16m-1)\sqrt{p}] \\ &= \frac{1}{4} [p - 32m\sqrt{p} + 3\sqrt{p}] \\ &\stackrel{m=\frac{p-1}{2d}}{=} \frac{1}{4} \left[ p - \frac{16(p-1)}{d}\sqrt{p} + 3\sqrt{p} \right] \\ &\stackrel{d \geq 16\sqrt{p-1}}{\geq} \frac{1}{4} \left[ p - \sqrt{(p-1)p} + 3\sqrt{p} \right] > \frac{3\sqrt{p}}{4} > 1. \end{aligned}$$

□

#### 4.3.2 Elliptic elements of high order

**Proposition 4.12.** *Assume  $Q_p$  is transitive and let  $x \in \mathbb{F}_p$  be elliptic with  $d = d_p(\pm x) \geq 16\sqrt{p+1}$ . If  $B \subsetneq Y^*(p)$  is a proper block of the action of  $Q_p$  on  $Y^*(p)$ , then  $B$  contains at most one solution with first coordinate  $[x]$ .*

The general proof strategy is the same as in Proposition 4.11, albeit with a few extra technical details. In the hyperbolic case, we used a parametrization of the elements of a cycle of  $\text{rot}_1$  as function over  $\mathbb{F}_p^*$ , which allowed us to use Weil's bound (Theorem 4.10 above). In the elliptic case, a similar approach requires that we go over the elements in the cyclic subgroup of size  $p+1$  in  $\mathbb{F}_{p^2}^*$ . The following lemma allows us to parametrize this subgroup as a function over  $\mathbb{F}_p$ :

**Lemma 4.13.** *The multiplicative subgroup  $H \leq \mathbb{F}_{p^2}^*$  of order  $p+1$  satisfies*

H

$$H = \left\{ \theta + i\eta \mid \theta, \eta \in \mathbb{F}_p, \theta^2 + \eta^2 = 1 \right\} = \left\{ \frac{2s}{1+s^2} + i\frac{1-s^2}{1+s^2} \mid s \in \mathbb{F}_p \right\} \cup \{-i\} \quad (10)$$

(where  $i = \sqrt{-1} \in \mathbb{F}_{p^2}$ ).

*Proof.* Note that  $(\theta + i\eta)^p = \theta - i\eta$  (recall that  $p \equiv 3(4)$  so  $i^p = i^{4k+3} = i^3 = -i$ ). So  $(\theta + i\eta)^{p+1} = (\theta + i\eta)(\theta - i\eta) = \theta^2 + \eta^2$ . This gives the first equality in (10). A straightforward computation yields the second equality. □

*Proof of Proposition 4.12.* We use the notation  $H$  for the subgroup of order  $p+1$  in  $\mathbb{F}_{p^2}^*$ , as in Lemma 4.13. Assume that  $[x, y_0, y_1]$  and  $[x, z_0, z_1]$  both belong to  $B$ . By Table 2,  $x = \omega + \omega^{-1}$  with  $\omega \in H$ , and we can assume that  $|\omega| = 2d \geq 32\sqrt{p+1}$ , for otherwise replace  $\omega$  by  $-\omega$  and  $x$  by  $-x$ . Write  $y_0 = A + A^p$ , with  $A \in \mathbb{F}_{p^2}$ , and  $A^{p+1} = \frac{x^2}{x^2-4}$ , and  $y_1 = A\omega + A^p\omega^{-1}$  (see Lemma 2.3). The cycle of  $\text{rot}_1$  containing  $[x, y_0, y_1]$  is

$$[x, y_0, y_1] = [x, y_d, y_{d+1}], [x, y_1, y_2], \dots, [x, y_{d-2}, y_{d-1}], [x, y_{d-1}, y_d]$$

with

$$y_j = A\omega^j + A^p\omega^{-j}.$$



The set  $\{\omega^j\}_{0 \leq j \leq 2d-1}$  is the same as the set  $\{h^m\}_{h \in H}$  where  $m = \frac{p+1}{2d}$ , with every element in  $\{\omega^j\}$  covered by  $m$  different values of  $h$ . So as sets,

$$\{y_0, \dots, y_{2d-1}\} = \{A\omega^j + A^p\omega^{-j}\}_{0 \leq j \leq 2d-1} = \left\{f_A(h) \stackrel{\text{def}}{=} Ah^m + A^p h^{-m}\right\}_{h \in H}.$$

The same holds for the cycle of  $\text{rot}_1$  containing  $[x, z_0, z_1]$  with  $C \in \mathbb{F}_{p^2}$  in the role of  $A$ , so that  $z_j = C\omega^j + C^p\omega^{-j}$ . We may assume that  $C \neq \pm A$ , for otherwise  $[x, y_0, y_1] = [x, z_0, z_1]$ . Moreover, if  $h^m = \omega^j$  then  $f_A(h) = y_j$  and  $f_C(h) = z_j$ .

As in the proof of Proposition 4.11, we derive from Lemma 4.6 that

$$\left(\frac{(f_A(h)^2 - 4)(f_C(h)^2 - 4)}{p}\right) = \left(\frac{(f_{A\omega}(h)^2 - 4)(f_{C\omega}(h)^2 - 4)}{p}\right) = -1 \quad (11)$$

cannot hold for any  $h \in H$ . To be able to use Theorem 4.10, we want to reparametrize (11) as polynomials in  $s \in \mathbb{F}_p$ , using Lemma 4.13. Denote

$$g_A(s) \stackrel{\text{def}}{=} (1 + s^2)^{2m} [f_A(h(s))^2 - 4]$$

where  $h(s) = \frac{2s+i(1-s^2)}{1+s^2} = \frac{-i(s+i)^2}{1+s^2}$ . Let also  $k_1 = g_A g_C$  and  $k_2 = g_{A\omega} g_{C\omega}$ . Then (11) is equivalent to

$$\left(\frac{k_1(s)}{p}\right) = \left(\frac{k_2(s)}{p}\right) = -1.$$

As in the proof of Proposition 4.11, for  $\ell \in \{\pm 1\}^2$ , denote by  $N_\ell$  the number of  $s \in \mathbb{F}_p$  for which  $\left(\frac{k_j(s)}{p}\right) = \ell_j$ ,  $j = 1, 2$ . Our goal is to get a contradiction by showing that  $N_{(-1, -1)} > 0$ . To this aim, for every  $B \subseteq \{1, 2\}$ , define

$$M_B \stackrel{\text{def}}{=} (-1)^{|B|} \sum_{s \in \mathbb{F}_p} \left(\frac{\prod_{j \in B} k_j(s)}{p}\right). \quad (12)$$

As in the proof of Proposition 4.11,  $g_A$  has no roots in  $\mathbb{F}_p$ , so

$$N_{(-1, -1)} = \frac{1}{4} \left( \sum_{B \subseteq \{1, 2\}} M_B \right). \quad (13)$$

We use Theorem 4.10 to estimate the  $M_B$ 's. First, we show that  $k_1, k_2 \in \mathbb{F}_p[x]$ . Notice that  $h(s)^{-1} = \frac{2s-i(1-s^2)}{1+s^2} = \frac{i(s-i)^2}{1+s^2}$ , so

$$\begin{aligned} g_A(s) &= (1 + s^2)^{2m} [f_A(h(s)) + 2][f_A(h(s)) - 2] \\ &= (1 + s^2)^{2m} (Ah(s)^m + A^p h(s)^{-m} + 2)(Ah(s)^m + A^p h(s)^{-m} - 2) \end{aligned} \quad (14)$$

$$\begin{aligned} &= \left(A[-i(s+i)^2]^m + A^p[i(s-i)^2]^m + 2[1+s^2]^m\right) \\ &\quad \cdot \left(A[-i(s+i)^2]^m + A^p[i(s-i)^2]^m - 2[1+s^2]^m\right). \end{aligned} \quad (15)$$

The last expression shows that  $g_A(s) \in \mathbb{F}_{p^2}[s]$ . Its degree is  $4m$ : indeed, the leading coefficient is

$$(-1)^m (A^2 + A^{2p}) + 2A^{p+1} - 4,$$

and for  $m$  even this coefficient equals  $(A + A^p)^2 - 4 = y_0^2 - 4$  which is not zero since  $y_0 \neq \pm 2$  (see Lemma 2.3). For  $m$  odd, this coefficient is

$$-(A + A^p)^2 + 4(A^{p+1} - 1) = -y_0^2 + 4(A^{p+1} - 1),$$

which is not zero because  $A^{p+1} - 1 = \frac{4}{x^2-4}$  is not a square in  $\mathbb{F}_p$  when  $x$  is elliptic.

As  $\mathbb{F}_{p^2} = \mathbb{F}_p + i\mathbb{F}_p$ , we can write  $g_A = g'_A + ig''_A$ , where  $g'_A, g''_A \in \mathbb{F}_p[s]$ . By definition, for every  $s \in \mathbb{F}_p$ , we have  $h = h(s) \in H$ , and

$$g_A(s) = (1 + s^2)^{2m} [f_A(h)^2 - 4] \in \mathbb{F}_p$$

so  $g''_A(s) = 0$  for every  $s \in \mathbb{F}_p$ . Since  $\deg(g''_A) \leq 4m < p$ , we conclude that  $g''_A$  is the zero polynomial, hence  $g_A(s) \in \mathbb{F}_p[s]$  and so  $k_1, k_2 \in \mathbb{F}_p[x]$ .

Next, we wish to show that  $k_1, k_2$  and  $k_1 k_2$  are not squares in  $\overline{\mathbb{F}_p}[x]$ . There is a one-to-one correspondence between the roots of  $g_A$  in  $\overline{\mathbb{F}_p}$  and the roots of

$$r_A(h) \stackrel{\text{def}}{=} (Ah^{2m} + 2h^m + A^p)(Ah^{2m} - 2h^m + A^p),$$

in  $\overline{\mathbb{F}_p}$  given by

$$\begin{aligned} \alpha &\mapsto h(\alpha) = \frac{-i(\alpha + i)^2}{1 + \alpha^2} \\ \frac{1 + ih}{h + i} &= \alpha(h) \leftrightarrow h, \end{aligned}$$

because  $\pm i$  is never a root of  $g_A$  and  $-i$  never a root of  $r_A$ . It is easier to analyze the roots of  $r_A$  than those of  $g_A$ : if  $h$  is a root of  $r_A$  then

$$h^m = \frac{\pm 1 \pm \sqrt{1 - \kappa(x)}}{A},$$

where  $\kappa(x) = A^{p+1} = \frac{x^2}{x^2-4}$ . Now note the following:

- The four possible values of  $h^m$  are distinct and different from zero (this follows from  $\kappa(x) \neq 0, 1$ ).
- Because  $(m, p) = 1$ , the four polynomials  $h^m - \frac{\pm 1 \pm \sqrt{1 - \kappa(x)}}{A}$  are separable, so  $r_A$  has  $4m$  distinct roots in  $\overline{\mathbb{F}_p}$ , and so does  $g_A$ .
- If  $A \neq \pm C$ , the  $4m$  roots of  $r_A$  are distinct from the  $4m$  roots of  $r_C$ : certainly  $\frac{1 + \sqrt{1 - \kappa(x)}}{A} \neq \pm \frac{1 + \sqrt{1 - \kappa(x)}}{C}$ , and if  $\frac{1 + \sqrt{1 - \kappa(x)}}{A} = \pm \frac{1 - \sqrt{1 - \kappa(x)}}{C}$  we obtain

$$\begin{aligned} C &= \pm A \cdot \frac{1 - \sqrt{1 - \kappa(x)}}{1 + \sqrt{1 - \kappa(x)}} \\ \kappa(x) = C^{p+1} &= A^{p+1} \left( \frac{1 - \sqrt{1 - \kappa(x)}}{1 + \sqrt{1 - \kappa(x)}} \right)^{p+1} = \kappa(x) \xi^{p+1} \end{aligned}$$

with  $\xi = \frac{1 - \sqrt{1 - \kappa(x)}}{1 + \sqrt{1 - \kappa(x)}} \in \mathbb{F}_p$  because  $1 - \kappa(x) = \frac{-4}{x^2-4}$  is a square in  $\mathbb{F}_p$ . Then  $\xi = \pm 1$ , that is,  $C = \pm A$  – a contradiction. Hence  $k_1 = g_A g_C$  and  $k_2 = g_{A\omega} g_{C\omega}$  are separable of degree  $8m$  each.

- Finally, if  $C \neq \pm A$ , the polynomial  $k_1 k_2 = g_A g_{A\omega} g_C g_{C\omega}$  is not a square in  $\overline{\mathbb{F}_p}[x]$ : it is separable unless  $A = \pm C\omega$  or  $A\omega = \pm C$ , but the two cannot hold simultaneously.

We can now apply Theorem 4.10 to obtain the same bounds on the  $M_B$ 's as in the proof of Proposition 4.11, and from (13) we now obtain

$$\begin{aligned}
N_{(-1,-1)} &\geq \frac{1}{4} [p - 2(8m-1)\sqrt{p} - (16m-1)\sqrt{p}] \\
&= \frac{1}{4} [p - 32m\sqrt{p} + 3\sqrt{p}] \\
&\stackrel{m=\frac{p+1}{2d}}{=} \frac{1}{4} \left[ p - \frac{16(p+1)}{d}\sqrt{p} + 3\sqrt{p} \right] \\
&\stackrel{d \geq 16\sqrt{p+1}}{\geq} \frac{\sqrt{p}}{4} [\sqrt{p} - \sqrt{p+1} + 3] > 0.
\end{aligned}$$

□

## 5 Strong Approximation for Square Free Composite Moduli

In this section we derive our main application for our results on the groups  $Q_p$  and show that  $\Gamma$  acts transitively on  $X^*(n)$  for various square-free composite values  $n = p_1 \cdots p_k$ . First, we prove Theorem 1.6 and show the transitivity under the assumption that  $Q_{p_j} \geq \text{Alt}(Y^*(p_j))$  for all  $j$ . Corollary 1.7 then follows using Theorems 1.3 and 1.4, the latter using the CFSG (so if all primes are 1 mod 4, this proof is also CFSG-free). Then, in Section 5.2, we give a slightly more involved proof to the transitivity of the  $\Gamma$ -action on  $X^*(n)$  under the milder assumption that  $Q_{p_j}$  is primitive for every  $j$ , thus proving Theorem 1.9 and deducing Corollary 1.7 without relying on the CFSG.

### 5.1 Transitivity modulo square free composite moduli, assuming alternating groups

Here Theorem 1.6 is proved. So  $n = p_1 \cdots p_k$  is a product of distinct primes, and for every  $1 \leq j \leq k$ ,  $Q_{p_j} \geq \text{Alt}(Y^*(p_j))$ . Our goal is to show that  $\Gamma_n$ , the permutation group induced by the action of  $\Gamma$  on  $X^*(n)$ , is transitive.

Our proof is by induction on  $k$ . The induction basis of  $k = 0$  being trivial. Denoting  $n' = p_1 \cdots p_{k-1}$ , assume that  $\Gamma$  acts transitively on  $X^*(n')$ . It is enough to show that  $\Lambda \stackrel{\text{def}}{=} \ker(\Gamma \rightarrow \Gamma_{n'})$  acts transitively on  $X^*(p_k)$ . To handle some special cases where the general argument does not work, we assume that if 2, 5 and or 11 are among the factors of  $n$ , then come first in the product  $p_1 \cdots p_k$  and in ascending order.

**Lemma 5.1.** *The action of  $\Lambda$  on  $Y^*(p_k)$  is transitive.*

*Proof.* We can assume<sup>7</sup>  $p_k \geq 5$ . Then  $|Y^*(p_k)| = \frac{p_k(p_k \pm 3)}{4} \geq 5$ , and so  $\text{Alt}(Y^*(p_k))$  is simple. This group is never a composition (Jordan-Hölder) factor of  $\Gamma_{p_j}$  if  $Q_{p_j} \geq \text{Alt}(Y^*(p_j))$  and  $p_j \neq p_k$ . To see it, note that the composition factors of  $\Gamma_{p_j}$  are either  $\text{Alt}(Y^*(p_j))$  (if  $p_j \geq 5$ ) or abelian, coming from  $Q_2 \cong \text{Sym}(4)$ , or from  $\text{Sym}(Y^*(p_j))/\text{Alt}(Y^*(p_j)) \cong \mathbb{Z}/2\mathbb{Z}$ , or from  $Q_p/\Gamma_p$  which is a subgroup of a power of  $\text{Sym}(4)$ .

Now consider the set

$$Z^*(n) \stackrel{\text{def}}{=} X^*(n') \times Y^*(p_k)$$

<sup>7</sup>For  $p = 2, 3$ , note that  $X^*(2) = Y^*(2)$  is of size 4 and  $\Gamma$  acts on it as  $\text{Sym}(4)$  (transitively, in particular), and  $X^*(3) = Y^*(3) = \emptyset$ .

and denote by  $T_n$  the permutation group on  $Z^*(n)$  induced by the action of  $\Gamma$ . By assumption,  $\text{Alt}(Y^*(p_k))$  is a composition factor of  $T_n$ . Consider now the normal sequence of subgroups

$$\ker(\Gamma \twoheadrightarrow T_n) \trianglelefteq \Lambda = \ker(\Gamma \twoheadrightarrow \Gamma_{p_1 \cdot p_2 \cdots p_{k-1}}) \trianglelefteq \cdots \trianglelefteq \ker(\Gamma \twoheadrightarrow \Gamma_{p_1 \cdot p_2}) \trianglelefteq \ker(\Gamma \twoheadrightarrow \Gamma_{p_1}) \trianglelefteq \Gamma.$$

The overall quotient is  $T_n$ . The quotient  $\ker(\Gamma \twoheadrightarrow \Gamma_{p_1 \cdots p_{j-1}}) / \ker(\Gamma \twoheadrightarrow \Gamma_{p_1 \cdots p_j})$  is a subgroup of  $\Gamma_j$ , so has no composition factor isomorphic to  $\text{Alt}(Y^*(p_k))$ . Thus the composition factor  $\text{Alt}(Y^*(p_k))$  must belong to the leftmost quotient  $\Lambda / \ker(\Gamma \twoheadrightarrow T_n)$ , which is exactly the permutation group induced by the action of  $\Lambda$  on  $Y^*(p_k)$ .  $\square$

The next step is to show that  $\Lambda$  acts transitively not only on  $Y^*(p_k)$  but also on  $X^*(p_k)$ . The general strategy is to find a triple  $(x, y, z) \in X^*(p_k)$  and elements in  $\Lambda$  mapping  $(x, y, z)$  to the other elements in its 4-block:  $(x, -y, -z)$ ,  $(-x, y, -z)$  and  $(-x, -y, z)$ . Together with the transitivity of  $\Lambda$  on  $Y^*(p_k)$ , this would complete the proof. As in other places in this paper, the argument in the case  $p_k \equiv 1(4)$  is simpler. For  $p_k = 2, 5, 11$  we verified the transitivity of  $\Gamma_{2 \cdot 5 \cdot 11}$  by computer.

**Lemma 5.2.** *If  $p_k \equiv 1(4)$  and  $p_k \geq 13$  then  $\Lambda$  acts transitively on  $X^*(p_k)$ .*

*Proof.* Take some  $x$  hyperbolic of maximal order (namely, the  $\text{rot}_1$ -cycles in  $C_1(x)$  are of length  $p-1 \geq 12$  each). Since 0 has order 4,  $x \neq 0$  and  $(0, x, ix) \in X^*(p_k)$ . Let  $(r, s, t) \in X^*(p_k)$  be another solution with  $r$  elliptic. As  $p \equiv 1(4)$ , there is a number  $t$  with  $t \equiv 2 \pmod{4}$  and  $t \equiv 0 \pmod{p+1}$ . Since all  $\text{rot}_1$ -cycles in  $C_1(0)$  have length 4, we get that  $\text{rot}_1^t$  fixes all four elements in  $[r, s, t]$  while mapping  $(0, x, ix) \mapsto (0, -x, -ix)$ . By Lemma 5.1, there is some  $g \in \Lambda$  mapping  $[0, x, ix] \mapsto [r, s, t]$ . The element  $h_1 = g^{-1} \cdot \text{rot}_1^{-t} \cdot g \cdot (\text{rot}_1^t)$  is in  $\Lambda$  and maps  $(0, x, ix) \mapsto (0, -x, -ix)$ .

Since  $x$  is maximal hyperbolic, its order is  $(p-1)$  which is divisible by 4 (here, the order is the order of the associated  $\omega$  – see Table 1). Hence  $-x$  is also maximal hyperbolic. Let now  $(r', s', t') \in X^*(p)$  be a solution with  $s'$  elliptic. Note that  $\frac{p^2-1}{4} \equiv 0 \pmod{p+1}$  while  $\frac{p^2-1}{2} \equiv \frac{p-1}{2} \pmod{p-1}$ . Thus  $\text{rot}_2^{(p^2-1)/4}$  fixes all four elements in  $[r', s', t']$  while mapping  $(0, x, ix) \mapsto (0, x, -ix)$  and  $(0, -x, -ix) \mapsto (0, -x, ix)$ . By Lemma 5.1, there is some  $g' \in \Lambda$  mapping  $[0, x, ix] \mapsto [r', s', t']$ . The element  $h_1 = (g')^{-1} \cdot \text{rot}_1^{-t} \cdot g' \cdot (\text{rot}_1^t)$  is in  $\Lambda$  and maps  $(0, x, ix) \mapsto (0, x, -ix)$  and  $(0, -x, -ix) \mapsto (0, -x, ix)$ .  $\square$

**Lemma 5.3.** *If  $p_k \equiv 3(4)$  and  $p_k \neq 3, 11$  then  $\Lambda$  acts transitively on  $X^*(p_k)$ .*

*Proof.* The argument is analogous to the one in the proof of Lemma 5.2: by Proposition 5.4 below, there is a solution  $(x, y, z) \in X^*(p_k)$  with both  $x$  and  $y$  elliptic of order divisible by 4. In this case,  $-x$  has the same order as  $x$ , say this order is  $4m \mid (p+1)$ . Let  $(r, s, t) \in X^*(p_k)$  be another solution with  $r$  hyperbolic. As  $p-1 \equiv 2(4)$ , there is a number  $t$  with  $t \equiv 2m \pmod{4m}$  and  $t \equiv 0 \pmod{p-1}$ . We get that  $\text{rot}_1^t$  fixes all four elements in  $[r, s, t]$  while mapping  $(x, y, z) \mapsto (x, -y, -z)$  and  $(-x, -y, z) \mapsto (-x, y, -z)$ . By Lemma 5.1, there is some  $g \in \Lambda$  mapping  $[x, y, z] \mapsto [r, s, t]$ . The element  $h_1 = g^{-1} \cdot \text{rot}_1^{-t} \cdot g \cdot (\text{rot}_1^t)$  is in  $\Lambda$  and maps  $(x, y, z) \mapsto (x, -y, -z)$  and  $(-x, -y, z) \mapsto (-x, y, -z)$ . In the same fashion, we find an element of  $\Lambda$  mapping  $(x, y, z) \mapsto (-x, y, -z)$  and we are done.  $\square$

**Proposition 5.4.** *For every prime  $p \neq 3, 11$  with  $p \equiv 3(4)$ , there is a solution  $(x, y, z) \in X^*(p)$  with two coordinates elliptic and order divisible by 4.*

Here we use notation as in Section 4.3.2. If  $\omega \in H$  is not a square and  $4 \mid (p+1)$ , then  $4 \mid |\omega|$ . Thus, it is enough to find a solution  $(x, y, z) \in X^*(p)$  with  $x, y$  elliptic and the corresponding  $\omega_x, \omega_y$  not squares in  $H$ .

**Lemma 5.5.** *Assume  $y = \omega + \omega^{-1}$  is elliptic (so  $\omega \in H$ ). Then  $\omega$  is a square in  $H$  if and only if  $y + 2$  is a square in  $\mathbb{F}_p$ .*

*Proof.* Note that  $y + 2 = \omega + \omega^{-1} + 2 = (\omega^{1/2} + \omega^{-1/2})^2$ . If  $\omega^{1/2} \in H$  then  $\omega^{1/2} + \omega^{-1/2} \in \mathbb{F}_p$ . On the other hand, if  $\omega^{1/2} \notin H$ , then  $\omega^{(p+1)/2} = -1$  and so  $\omega^{1/2} + \omega^{-1/2} \notin \mathbb{F}_p$ , because

$$\left(\omega^{1/2} + \omega^{-1/2}\right)^p = \omega^{(p+1)/2} \omega^{-1/2} + \omega^{-(p+1)/2} \omega^{1/2} = -\left(\omega^{-1/2} + \omega^{1/2}\right) \neq \left(\omega^{1/2} + \omega^{-1/2}\right)$$

(the last inequality stems from  $(\omega^{1/2} + \omega^{-1/2})^2 = y + 2 \neq 0$ ).  $\square$

*Proof of Proposition 5.4.* Fix  $x \in \mathbb{F}_p$  elliptic of maximal order  $(p+1)$ . So  $4|\omega_x| = p+1$ . By Lemma 5.5, it is enough to find  $y, z \in \mathbb{F}_p$  such that  $(x, y, z) \in X^*(p)$  is a solution,  $y$  is elliptic and  $y + 2$  is a non-square. Since  $y$  elliptic means that  $y^2 - 4 = (y+2)(y-2)$  is not a square, we need to find  $y, z$  with  $(x, y, z) \in X^*(p)$  and  $y + 2$  a non-square and  $y - 2$  a square.

Imitating the notation of the proof of Proposition 4.12, assume  $x = \omega + \omega^{-1}$  with  $\omega \in H$ , choose some  $A \in \mathbb{F}_{p^2}$  for which  $A^{p+1} = \frac{x^2}{x^2-4}$ , and let  $f_A(h) = Ah + A^p h^{-1}$  for  $h \in H$ . Then,

$$\{(f_A(h), f_{A\omega}(h)) \mid h \in H\} = \{(y, z) \mid (x, y, z) \in X^*(p)\}. \quad (16)$$

Recall the parametrization of  $H \setminus \{-i\}$  by elements from  $\mathbb{F}_p$  described in Lemma 4.13:  $h(s) = \frac{2s+i(1-s^2)}{1+s^2}$ . Define  $g_1, g_2 \in \mathbb{F}_p[s]$  as follows:

$$\begin{aligned} g_1(s) &\stackrel{\text{def}}{=} (1+s^2)^2 [f_A(h(s)) + 2] = (1+s^2) [2s(A + A^p) + (1-s^2)i(A - A^p) + 2(1+s^2)] \\ g_2(s) &\stackrel{\text{def}}{=} (1+s^2)^2 [f_A(h(s)) - 2] = (1+s^2) [2s(A + A^p) + (1-s^2)i(A - A^p) - 2(1+s^2)]. \end{aligned}$$

It is not hard to see that  $g_j(s) \in \mathbb{F}_p[s]$ : indeed,  $A + A^p, i(A - A^p) \in \mathbb{F}_p$ . We now show that for large enough  $p$ , there is some  $s \in \mathbb{F}_p$  for which

$$\left(\frac{g_1(s)}{p}\right) = -1 \quad \text{and} \quad \left(\frac{g_2(s)}{p}\right) = 1. \quad (17)$$

For  $\ell \in \{\pm 1\}^2$ , denote by  $N_\ell$  the number of  $s \in \mathbb{F}_p$  for which  $\left(\frac{g_j(s)}{p}\right) = \ell_j$  for  $j = 1, 2$ . We show that for large enough  $p$  we have  $N_{(-1,1)} > 0$ . For every set of coordinates  $B \subseteq \{1, 2\}$ , let  $M_B$  denote the difference between the number of  $s \in \mathbb{F}_p$  for which  $\left(\frac{\prod_{j \in B} g_j(s)}{p}\right)$  has the same sign as for a solution of (17) and the number of these with opposite sign, namely

$$M_B \stackrel{\text{def}}{=} (-1)^{\mathbf{1}_{1 \in B}} \sum_{s \in \mathbb{F}_p} \left(\frac{\prod_{j \in B} g_j(s)}{p}\right).$$

Consider the sum  $\sum_{B \subseteq \{1, 2\}} M_B$ . Every  $s \in \mathbb{F}_p$  which is a solution to (17) contributes 4 to this sum. Every other  $s \in \mathbb{F}_p$  contributes exactly 0, because the  $g_j$ 's have no roots in  $\mathbb{F}_p$ . For example, if  $g_1(s) = 0$  for some  $s \in \mathbb{F}_p$ , then  $f_A(h(s)) = -2$ , but  $X^*(p)$  has no solutions involving  $\pm 2$  when  $p \equiv 3(4)$ . Thus,

$$N_{(-1,1)} = \frac{1}{4} \left( \sum_{B \subseteq \{1, 2\}} M_B \right). \quad (18)$$

Of course,  $M_\emptyset = p$ . The argument in Proposition 4.12 shows that except for  $\pm i$ , the other roots of the  $g_j$ 's are all distinct from each other. Thus, none of  $g_1, g_2$  or  $g_1 g_2$  is a square in

$\overline{\mathbb{F}}_p[x]$ . Now  $g_1$  and  $g_2$  have each at most 4 distinct roots and by Theorem 4.10,  $M_B \geq -3\sqrt{p}$ . Their product  $g_1g_2$  has at most 6 distinct roots, hence by Theorem 4.10  $M_B \geq -5\sqrt{p}$ . From (18) we get

$$N_{(-1,1)} \geq \frac{1}{4}(p - 2 \cdot 3\sqrt{p} - 5\sqrt{p}) = \frac{p - 11\sqrt{p}}{4}.$$

So for  $p > 11^2 = 121$  we have  $N_{(-1,1)} > 0$  and we are done.

For all primes  $p$  with  $p \equiv 3(4)$ ,  $p \leq 121$  and  $p \neq 3, 11$ , we verified by a computer there is a solution  $(x, y, z) \in X^*(p)$  with  $x, y$  elliptic and of order divisible by 4. For example, one can take  $(3, 3, 3) \in X^*(7)$ ,  $(6, 6, 8) \in X^*(19)$ ,  $(3, 3, 3) \in X^*(23)$  and  $(4, 4, 9) \in X^*(31)$ .  $\square$

## 5.2 Transitivity modulo square free composite moduli, assuming primitivity

Here Theorem 1.9 is proved without the CFSG. Recall that Theorem 1.9 states that if  $n = p_1 \cdots p_k$  is a product of distinct primes and  $Q_{p_j}$  is a primitive permutation group for all  $j = 1, \dots, k$ , then  $\Gamma_n$  acts transitively on  $X^*(n)$ . Note that in Section 5.1, the only time we use the assumption on  $Q_{p_j}$  is in Lemma 5.1. Here, then, we prove an analogue to this lemma which relies only on the primitivity of the  $Q_{p_j}$ 's, and then apply Lemmas 5.2 and 5.3.

The main ingredients are the following theorems from the theory of finite permutation groups. The proofs are elementary and rather short, sometimes even several lines long, can all be found in [DM96], mostly in Section 4.3.

**Theorem 5.6** (See [DM96, Theorem 1.6A]). *If  $G \leq \text{Sym}(n)$  is a primitive permutation group and  $1 \neq H \trianglelefteq G$  is a non-trivial normal subgroup, then  $H$  is transitive.*

A minimal normal subgroup of a non-trivial group  $G$  is a normal subgroup  $K \neq 1$  of  $G$  which does not contain properly any other non-trivial normal subgroup of  $G$ . The *socle* of  $G$ , denoted  $\text{soc}(G)$ , is the subgroup generated by the set of all minimal normal subgroups of  $G$ . Note that the set of minimal normal subgroups is closed under conjugation, so  $\text{soc}(G) \trianglelefteq G$ .

**Theorem 5.7** (See [DM96, Theorems 4.3B, Corollary 4.3B and Theorem 4.7A]). *Let  $G \leq \text{Sym}(n)$  be a primitive subgroup. Then exactly one of the following holds:*

1. *For some prime  $p$  and some integer  $d$ , the group  $G$  is permutation isomorphic<sup>8</sup> to a subgroup of the affine group  $\text{Aff}(p, d)$  acting on  $\mathbb{F}_p^d$ , so, in particular,  $n = p^d$ . In this case,  $\text{soc}(G)$  is a regular<sup>9</sup> elementary abelian subgroup of order  $p^d$ .*
2.  *$\text{soc}(G) = K_1 \times K_2$  where  $K_1, K_2 \trianglelefteq G$  are minimal normal subgroups of  $G$ , which are regular, non-abelian and permutation isomorphic to each other. Moreover<sup>10</sup>,  $C_G(K_1) = K_2$  and  $C_G(K_2) = K_1$ . In addition,  $K_1 \cong K_2 \cong T^m$  for some finite simple non-abelian group  $T$  and some  $m$ .*
3.  *$\text{soc}(G)$  is a minimal normal subgroup of  $G$  and is non-abelian. Moreover,  $C_G(\text{soc}(G)) = 1$  and  $\text{soc}(G) \cong T^m$  for some finite simple non-abelian group  $T$  and some  $m$ .*

**Lemma 5.8.** *Assume  $q < p$  are distinct primes with  $Q_p$  primitive. Then  $\text{soc}(Q_p)$  is contained in the image of  $\ker(\Gamma \twoheadrightarrow \Gamma_q)$  in  $Q_p$ .*

<sup>8</sup>Two permutation groups are permutation isomorphic if they are the same permutation groups except for, possibly, the labeling of the points in the sets they act on.

<sup>9</sup>A permutation group  $H \leq \text{Sym}(n)$  is called *regular* if it is sharply transitive. Namely, it is transitive and free. In other words, it is transitive and of order  $n$ .

<sup>10</sup>For  $G$  a group and  $K \leq G$  a subgroup,  $C_G(K) = \{g \in G \mid gk = kg \ \forall k \in K\}$  is the centralizer of  $K$  in  $G$ .



*Proof.* Consider the socle subgroup  $\text{soc}(Q_p)$  of the primitive group  $Q_p$ . Case (1) of Theorem 5.7 is ruled out because  $|Y^*(p)|$  is not a prime power (or, alternatively, because  $\text{Aff}(p, d)$  has no non-identity elements fixing more than half of the points, such as  $\text{rot}_1^{p(p+1)/2} \in Q_p$ ). So either  $Q_p$  belongs to case (2) or it belongs to case (3).

Denote by  $\pi_p: \Gamma \rightarrow Q_p$  the projection. We show first that  $\text{soc}(Q_p) \leq \pi_p(\ker(\Gamma \rightarrow Q_q))$ . The order of  $\text{rot}_1$  in  $Q_p$  is  $\frac{p^2-1}{4}$ . Therefore, the element  $g = \text{rot}_1^{(q^2-1)/4}$  is trivial in  $Q_q$  but not in  $Q_p$  (by abuse of notation, we regard  $g$  as an element of  $\Gamma$  and of its quotients  $Q_q$  and  $Q_p$ ). Assume first that  $Q_p$  belongs to case (3). Since  $C_{Q_p}(\text{soc}(Q_p)) = 1$ , there is some  $h \in \text{soc}(Q_p)$  not commuting with  $g \in Q_p$ , so  $e \neq [g, h] = ghg^{-1}h^{-1} \in \text{soc}(Q_p) \cap \pi_p(\ker(\Gamma \rightarrow Q_q))$ . Since  $\text{soc}(Q_p)$  is a minimal normal subgroup of  $Q_p$ , it is generated by the conjugates of  $[g, h]$  in  $Q_p$ , all of which also belong to  $\pi_p(\ker(\Gamma \rightarrow Q_q))$ . Thus  $\text{soc}(Q_p) \leq \pi_p(\ker(\Gamma \rightarrow Q_q))$ .

Now assume that  $Q_p$  belongs to case (2). Since regular subgroups of  $\text{Sym}(n)$  are obtained as the (left or right) regular representation of a group of order  $n$ , every element of a regular permutation group has all its cycles with equal length. Since  $\text{rot}_1 \in Q_p$  contains cycles of coprime lengths, no non-trivial power of it can belong to a regular subgroup, so  $g = \text{rot}_1^{(q^2-1)/4} \notin K_1 \cup K_2$ . So there are  $h_1 \in K_1$  and  $h_2 \in K_2$  not commuting with  $g$ . Consider  $h = h_1h_2 \in K_1 \times K_2 = \text{soc}(Q_p)$ . Then  $[g, h] = ([g, h_1], [g, h_2]) \in K_1 \times K_2 = \text{soc}(Q_p)$  belongs also to  $\pi_p(\ker(\Gamma \rightarrow Q_q))$  but not to  $K_1 \cup K_2$ . The only normal subgroups of  $Q_p$  which are contained in  $K_1 \times K_2$  are  $1, K_1, K_2$  and  $K_1 \times K_2$ . Hence  $K_1 \times K_2$  is generated by the conjugates in  $Q_p$  of  $[g, h]$ , all of which belong to  $\pi_p(\ker(\Gamma \rightarrow Q_q))$ . Thus  $\text{soc}(Q_p) \leq \pi_p(\ker(\Gamma \rightarrow Q_q))$ .

Finally, the quotient  $\ker(\Gamma \rightarrow Q_q)/\ker(\Gamma \rightarrow \Gamma_q)$  is isomorphic to the subgroup of  $\Gamma_q$  fixing the 4-blocks, so it is a subgroup of  $\text{Sym}(4)^{|Y^*(p)|}$ . In particular, it is solvable and so is  $\pi_p(\ker(\Gamma \rightarrow Q_q))/\pi_p(\ker(\Gamma \rightarrow \Gamma_q))$ . So inside  $Q_p$ , some derived group of  $\pi_p(\ker(\Gamma \rightarrow Q_q))$  is contained in  $\pi_p(\ker(\Gamma \rightarrow \Gamma_q))$ . By Theorem 5.7,  $\text{soc}(Q_p) \cong T^m$  for some finite simple non-abelian group  $T$  and some  $m$ , whence  $\text{soc}(Q_p)$  is perfect, namely,  $[\text{soc}(Q_p), \text{soc}(Q_p)] = \text{soc}(Q_p)$ . This yields that  $\text{soc}(Q_p)$  is contained in any derived subgroup<sup>11</sup> of  $\pi_p(\ker(\Gamma \rightarrow \Gamma_q))$ . Thus  $\text{soc}(Q_p) \leq \pi_p(\ker(\Gamma \rightarrow \Gamma_q))$ .  $\square$

*Proof of Theorem 1.9.* Assume  $n = p_1 \cdots p_k$  is a product of distinct primes with  $Q_{p_1}, \dots, Q_{p_k}$  primitive and  $p_1 < p_2 < \dots < p_k$ . We prove that  $\Gamma$  acts transitively on  $X^*(n)$  by induction on  $k$ , with the trivial induction basis of  $k = 0$ . We may also assume that  $p_k \geq 13$ : for  $p \leq 11$ , we verified by computer that  $Q_p = \text{Sym}(Y^*(p))$ , so Theorem 1.6 applies.

Assume that  $\Gamma_{p_1 \cdots p_{k-1}}$  is transitive. If  $\pi_{p_k}: \Gamma \rightarrow Q_{p_k}$  is the projection and  $\Lambda \stackrel{\text{def}}{=} \ker(\Gamma \rightarrow \Gamma_{p_1 \cdots p_{k-1}})$ , then Lemma 5.8 guarantees that

$$\text{soc}(Q_p) \leq \pi_p(\ker(\Gamma \rightarrow \Gamma_{p_1})) \cap \dots \cap \pi_p(\ker(\Gamma \rightarrow \Gamma_{p_{k-1}})) = \pi_p(\ker(\Gamma \rightarrow \Gamma_{p_1 \cdots p_{k-1}})) = \pi_p(\Lambda).$$

Since  $\text{soc}(Q_p)$  is a normal subgroup of the primitive group  $Q_p$ , it is transitive by Theorem 5.6. Hence  $\Lambda$  acts transitively on  $Y^*(p_k)$ . This proves the analogue of Lemma 5.1. The rest of the proof is as in Section 5.1: note that the proofs of Lemmas 5.2 and 5.3 do not use the assumption of  $Q_p$  containing the alternating group.  $\square$

## 6 $T_2$ -systems

This section explains why Theorem 1.11 is equivalent to Theorems 1.3 and 1.4. Namely, if we let  $\Sigma_{2,-2}(p)$  denote the set of  $\text{PSL}(2, p)$ -defining subgroups of  $F_2$  with associated trace  $-2$ , our goal here is to show:

<sup>11</sup>By a derived subgroup of the group  $G$ , we mean one of  $G^{(i)}$ , where  $G^{(0)} = G$  and  $G^{(i)} = [G^{(i-1)}, G^{(i-1)}]$ .

1. A one-to-one correspondence between  $Y^*(p)$  and  $\Sigma_{2,-2}(p)$ , and
2. An isomorphism between  $Q_p$ , the group of permutations induced by the action of  $\Gamma$  on  $Y^*(p)$ , and the group of permutations induced by the action of  $\text{Aut}(\mathbb{F}_2)$  on  $\Sigma_{2,-2}(p)$ .

First, let us define  $\Sigma_{2,-2}(p)$  properly. For  $A, B \in \text{PSL}(2, p)$ , define

$$\text{Tr}(A, B) \stackrel{\text{def}}{=} (\text{tr}A, \text{tr}B, \text{tr}AB) \in \mathbb{F}_p^3 / \sim, \quad (19)$$

where  $\sim$  is the equivalence of changing the sign of two of the coordinates (each of  $A$  and  $B$  is a well-defined matrix in  $\text{SL}(2, p)$  up to a sign). Assume  $\langle A, B \rangle = \text{PSL}(2, p)$ , and let  $\varphi: \mathbb{F}_2 \twoheadrightarrow \text{PSL}(2, p)$  be the epimorphism mapping the generators  $a$  and  $b$  of  $\mathbb{F}_2$  to  $A$  and  $B$ , respectively. The kernel  $N = \ker \varphi$  is a  $\text{PSL}(2, p)$ -defining subgroup of  $\mathbb{F}_2$ , and define

$$\text{Tr}(N) \stackrel{\text{def}}{=} \text{Tr}(A, B).$$

Recall that  $\Sigma_2(G)$  denotes the set of  $G$ -defining subgroups of  $\mathbb{F}_2$ .

*Claim 6.1.* The map  $\text{Tr}: \Sigma_2(\text{PSL}(2, p)) \rightarrow \mathbb{F}_p^3 / \sim$  is well-defined.

*Proof.* Let  $G = \text{PSL}(2, p)$ . Given  $N \in \Sigma_2(G)$ , all epimorphisms  $\mathbb{F}_2 \twoheadrightarrow G$  with kernel  $N$  are obtained one from the other by post-composition with some automorphism from  $\text{Aut}(G)$ . But every automorphism of  $G$  is obtained by a conjugation by some element from  $\text{PGL}(2, p)$ . Evidently, such conjugation does not effect the image of  $\text{Tr}$  on the images of the generators  $a$  and  $b$  of  $\mathbb{F}_2$ .  $\square$

Recall that  $\text{tr}([A, B]) = Q(\text{tr}A, \text{tr}B, \text{tr}AB)$  where  $Q(x, y, z) = x^2 + y^2 + z^2 - xyz - 2$ . Thus, for  $N \in \Sigma_2(\text{PSL}(2, p))$ , the element

$$Q(N) \stackrel{\text{def}}{=} Q(\text{Tr}(N)) \in \mathbb{F}_p$$

is well-defined, and we denote

$$\Sigma_{2,-2}(p)$$

$$\Sigma_{2,-2}(p) \stackrel{\text{def}}{=} Q^{-1}(-2) \subseteq \Sigma_2(\text{PSL}(2, p)).$$

Note that, by definition, for every  $N \in \Sigma_{2,-2}(p)$  the triple  $\text{Tr}(N)$  is (an equivalence class up to sign changes of) a solution to the Markoff equation (1) over  $\mathbb{Z}/p\mathbb{Z}$ .

*Claim 6.2.* The map  $\text{Tr}|_{\Sigma_{2,-2}(p)}$  is a bijection from  $\Sigma_{2,-2}(p)$  to  $Y^*(p)$ .

*Proof.* Consider the map  $\widetilde{\text{Tr}}: \text{SL}(2, p) \times \text{SL}(2, p) \rightarrow \mathbb{F}_p^3$  defined as in (19). By [Mac69, Theorems 2 and 3], if  $(x, y, z) \in \mathbb{F}_p^3$  is the image of some generating pair in  $\text{SL}(2, p)$ , then every two pairs in  $\widetilde{\text{Tr}}^{-1}((x, y, z))$  are conjugated one to the other by an element  $g \in \text{SL}(2, \overline{\mathbb{F}_p})$ . Since these pairs are generating, this conjugation by  $g$  is an automorphism of  $\text{SL}(2, p)$ . As every automorphism of  $\text{SL}(2, p)$  is also an automorphism of  $\text{PSL}(2, p)$ , we obtain that

$$\text{Tr}|_{\Sigma_{2,-2}(p)}: \Sigma_{2,-2}(p) \rightarrow \mathbb{F}_p^3 / \sim$$

is injective.

By [Mac69, Thm 1], the map  $\widetilde{\text{Tr}}$  is surjective. The analysis in [MW13, Section 11] shows that the only triple  $(x, y, z) \in \mathbb{F}_p^3$  with  $Q(x, y, z) = -2$  which does not correspond to generating pairs is<sup>12</sup>  $(0, 0, 0)$ . This completes the proof of the claim.  $\square$

<sup>12</sup>To see that  $(0, 0, 0)$  is not associated with a generating pair, note that if  $A \in \text{PSL}(2, p)$  has trace 0, then  $A$  is an involution. If both  $A$  and  $B$  are involutions, then  $\langle A, B \rangle$  is a dihedral group, which is a proper subgroup of  $\text{PSL}(2, p)$ .

We have left to show the isomorphism of  $Q_p$  and the permutation group induced by  $\text{Aut}(F_2)$  on  $Y^*(p) \cong \Sigma_{2,-2}(p)$ . Recall that  $Q_p = \langle \tau_{(12)}, \tau_{(23)}, R_3 \rangle$ . For  $F_2 = F(a, b)$ ,  $\text{Aut}(F_2)$  is generated by the following Nielsen moves<sup>13</sup>:  $r: (a, b) \mapsto (a^{-1}, b)$ ,  $s: (a, b) \mapsto (b, a)$  and  $t: (a, b) \mapsto (a^{-1}, ab)$ . The induced action of these three automorphisms on  $Y^*(p)$  is easily seen to be the same action given by  $R_3$ ,  $\tau_{(12)}$  and  $\tau_{(23)}$ , respectively.

## Appendix

### A On the order of a quadratic integer modulo most primes By Dan Carmon

Throughout this appendix, we use the notation  $f \ll g$  to mean that there exists an absolute constant  $C > 0$  for which  $f \leq Cg$  for all valid values of the implicit variables. The similar notation  $f \ll_a g$  means there exists a function  $C = C(a) > 0$  for which  $f \leq Cg$ . The notation  $f \asymp g$  is shorthand for “ $f \ll g$  and  $g \ll f$ ”.

#### The main claim

Let  $a \in \mathbb{Q}(\sqrt{D})$  be a fixed quadratic integer with norm 1 and absolute value  $|a| > 1$  (e.g.  $a = \frac{3+\sqrt{5}}{2}$ ). For primes  $p \nmid D$ , consider the residue  $\bar{a} = (a \bmod p)$ , as an element of either  $\mathbb{F}_p$  or  $\mathbb{F}_{p^2}$ , depending on whether  $D$  is a quadratic residue modulo  $p$ . In both cases there are two possible choices for  $\bar{a}$ , but its order  $o_p(a)$ , which is the smallest positive integer satisfying  $\bar{a}^{o_p(a)} = 1 \in \mathbb{F}_{p^2}$  is well-defined. Let  $\pi(x) = \#\{p \leq x : p \text{ prime}\}$  be the prime counting function. We prove the following claim:

**Proposition A.1.** *For any constant  $C \geq 1$ ,*

$$\#\{p \leq x : o_p(a) \leq C\sqrt{x}\} \ll_a \frac{\pi(x)}{(\log x)^\delta (\log \log x)^{3/2-\delta}}, \quad (20)$$

where  $\delta$  is the Erdős-Tenenbaum-Ford constant,

$$\delta = 1 - \frac{1 + \log \log 2}{\log 2} = 0.086071\dots$$

In particular, the set of primes with  $o_p(a) > C\sqrt{p}$  has relative density 1.

#### Proof outline

Proposition A.1 follows from the combination of two sub-lemmas:

**Lemma A.2.** *Let  $\alpha = \alpha(x)$  tend to infinity arbitrarily slowly with  $x$ , and let  $y = \sqrt{\frac{x}{\alpha}}$ . Then*

$$\#\{p \leq x : o_p(a) \leq y\} \ll_a \frac{\pi(x)}{\alpha}. \quad (21)$$

**Lemma A.3.** *Let  $\alpha, y$  be as in the previous Lemma. Define  $z = C\sqrt{x}$ , and  $u_0 = \frac{\log \alpha}{\log x}$ . Suppose further that  $\alpha \in \left(\frac{4}{C^2}, \frac{\sqrt{x}}{C}\right)$ . Then*

$$\#\{p \leq x : \exists d \in (y, z], p \equiv \pm 1 \pmod{d}\} \ll u_0^\delta \left(\log \frac{2}{u_0}\right)^{-3/2} \pi(x). \quad (22)$$

---

<sup>13</sup>We deliberately copy the notation for these Nielsen moves from [MW13].

Indeed, since  $a$  has norm 1,  $o_p(a)$  is always a factor of either  $p - 1$  when  $D$  is a quadratic residue modulo  $p$ , or of  $p + 1$  when  $D$  is a non-quadratic residue, i.e.  $p \equiv \pm 1 \pmod{o_p(a)}$  in either case. Thus  $o_p(a) \leq C\sqrt{x}$  implies that  $p$  is either included in the set of the first lemma if  $o_p(a) \leq y$ , or in the set of the second lemma if  $o_p(a) \in (y, z]$ . Choosing the optimal value

$$\alpha = (\log x)^\delta (\log \log x)^{3/2-\delta} \quad (23)$$

yields the claimed value in the right hand side of both lemmas.

## Proofs of the lemmas

*Proof of Lemma A.2.* The following proof is an adaptation of an argument from Erdős and Murty [EM99, Introduction], in which only integral values  $a$  and a specific choice of  $\alpha$  were considered.

For every  $k \geq 1$  define  $A_k = \frac{a^k - a^{-k}}{\sqrt{D}}$ . Note that  $A_k$  is always an integer, with  $|A_k| < |a|^k$ , and that  $o_p(a) = k$  implies  $p \mid A_k$ . Define

$$B_y = \prod_{k=1}^{\lfloor y \rfloor} A_k,$$

so that  $o_p(a) \leq y$  implies  $p \mid B_y$ . We now observe that

$$\log B_y = \sum_{k=1}^{\lfloor y \rfloor} \log A_k \leq \sum_{k=1}^{\lfloor y \rfloor} k \log |a| \ll_a y^2 = \frac{x}{\alpha}, \quad (24)$$

and on the other hand

$$\begin{aligned} \log B_y &\geq \sum_{p \mid B_y} \log p \geq \sum_{p : o_p(a) \leq y} \log p \geq \sum_{\substack{\sqrt{x} < p \leq x \\ o_p(a) \leq y}} \log \sqrt{x} \\ &= \frac{1}{2} \log x \cdot \#\{\sqrt{x} < p \leq x : o_p(a) \leq y\}, \end{aligned} \quad (25)$$

whence

$$\#\{p \leq x : o_p(a) \leq y\} \leq \pi(\sqrt{x}) + \frac{2 \log B_y}{\log x} \ll_a \frac{2}{\alpha} \frac{x}{\log x} \ll \frac{\pi(x)}{\alpha}. \quad (26)$$

□

*Proof of Lemma A.3.* This lemma is a direct application of results due to Ford [For08]. We cite the relevant definitions and theorems. Ford's main object of study is the function

$$H(x, y, z) = \#\{n \leq x : \exists d \in (y, z], d \mid n\}.$$

We are particularly interested in the specialized function

$$H(x, y, z; P_\lambda) = \#\{n \leq x : n \in P_\lambda, \exists d \in (y, z], d \mid n\},$$

where  $P_\lambda = \{p + \lambda : p - \text{prime}\}$  is a set of shifted primes, and more specifically only for  $\lambda = \pm 1$ .

In [For08, Theorem 1], Ford estimates  $H(x, y, z)$  for all possible choices of  $y \leq z \leq x$ . The relevant case for our choice of  $y, z$  is the third subcase of case (v), wherein  $x, y, z$  are all large,

$y \leq \sqrt{x}$ , and  $z \in [2y, y^2]$ , all of which are immediately validated for our values, due to the constraint on  $\alpha$ . For this case, the theorem states

$$\frac{H(x, y, z)}{x} \asymp u^\delta \left(\log \frac{z}{u}\right)^{-3/2}, \quad (27)$$

where  $u$  is the number satisfying  $z = y^{1+u}$ , or equivalently

$$u = \frac{\log(z/y)}{\log y} = \frac{\log(C\sqrt{\alpha})}{\log(\sqrt{x/\alpha})} = \frac{\log \alpha + 2 \log C}{\log x - \log \alpha} \asymp \frac{\log \alpha}{\log x} = u_0. \quad (28)$$

In [For08, Theorem 6], Ford estimates  $H(x, y, z; P_\lambda)$ , for any fixed non-zero  $\lambda$ . The behaviour of the function is determined by whether  $z$  is greater or lesser than  $y + (\log y)^{2/3}$ . The constraint on  $\alpha$  implies  $z \geq 2y$ , so we are certainly in the regime of  $z \geq y + (\log y)^{2/3}$ , in which the theorem yields

$$H(x, y, z; P_\lambda) \ll_\lambda \frac{H(x, y, z)}{\log x}. \quad (29)$$

Combining the estimates (27), (28), (29) yields (22), proving the lemma.  $\square$

## References

- [BGS16] Jean Bourgain, Alexander Gamburd, and Peter Sarnak. Markoff triples and strong approximation. *Comptes Rendus Mathématique*, 354(2):131–135, 2016.
- [BGS17] Jean Bourgain, Alexander Gamburd, and Peter Sarnak. Markoff surfaces and strong approximation: 1. arXiv preprint arXiv:1607.01530, 2017+.
- [CGMP16] Alois Cerbu, Elijah Gunther, Michael Magee, and Luke Peilen. The cycle structure of a Markoff automorphism over finite fields. preprint arXiv:1610.07077, 2016.
- [DM96] John D. Dixon and Brian Mortimer. *Permutation groups*. Springer Science & Business Media, 1996.
- [EM99] Pál Erdős and M. Ram Murty. On the order of a mod  $p$ . *CRM Proceedings and Lecture Notes*, 19:87–97, 1999.
- [For08] Kevin Ford. The distribution of integers with a divisor in a given interval. *Annals of mathematics*, pages 367–433, 2008.
- [Gil77] Robert Gilman. Finite quotients of the automorphism group of a free group. *Canad. J. Math*, 29(3):541–551, 1977.
- [GM98] Robert Guralnick and Kay Magaard. On the minimal degree of a primitive permutation group. *Journal of Algebra*, 207(1):127–145, 1998.
- [GS09] Shelly Garion and Aner Shalev. Commutator maps, measure preservation, and T-systems. *Transactions of the American Mathematical Society*, 361(9):4631–4651, 2009.
- [Lub11] Alexander Lubotzky. Dynamics of  $\text{Aut}(\text{Fn})$  actions on group presentations and representations. In B. Farb and D. Fisher, editors, *Geometry, Rigidity, and Group Actions*, pages 609–643. Chicago University press, 2011.
- [Mac69] Alexander M. Macbeath. Generators of the linear fractional groups. In *Proc. Symp. Pure Math*, volume 12, pages 14–32, 1969.

- [Mar79] Andrey Markoff. Sur les formes quadratiques binaires indéfinies. *Mathematische Annalen*, 15(3):381–406, 1879.
- [Mar80] Andrey Markoff. Sur les formes quadratiques binaires indéfinies. *Mathematische Annalen*, 17(3):379–399, 1880.
- [MW13] Darryl McCullough and Marcus Wanderley. Nielsen equivalence of generating pairs of  $SL(2, q)$ . *Glasgow Mathematical Journal*, 55(03):481–509, 2013.
- [Pak01] Igor Pak. What do we know about the product replacement algorithm? In W. Kantor and A. Seress, editors, *Groups and Computation III*, pages 301–347. de Gruyter, 2001.
- [Sch76] Wolfgang M. Schmidt. *Equations over finite fields: An elementary approach*. Springer-Verlag, Halsted Press, 1976.

Chen Meiri,  
 Department of Mathematics,  
 Technion - Israel Institute of Technology  
 Haifa 32000 Israel  
 chenm@tx.technion.ac.il

Doron Puder,  
 School of Mathematical Sciences,  
 Tel-Aviv University,  
 Tel-Aviv 69978 Israel  
 doronpuder@gmail.com

Dan Carmon,  
 School of Mathematical Sciences,  
 Tel-Aviv University,  
 Tel-Aviv 69978 Israel  
 dancarmo@post.tau.ac.il